

Konstruktionen im Lexikon

Konstruktionen in der Erzählanalyse

Abhandlung
zur Erlangung der Doktorwürde
der
Philosophischen Fakultät
der
Universität Zürich

Vorgelegt von
Marc Luder
von Höchstetten / BE

Angenommen im Frühjahrssemester 2011
auf Antrag von Frau Prof. Dr. Brigitte Boothe und
Herrn Prof. Dr. Arnulf Deppermann

Zürich, 2011

Dank

Ich möchte mich bei allen Personen ganz herzlich bedanken, die mich während der Entstehung dieser Arbeit beraten und unterstützt haben. Mein Dank geht in erster Linie an Frau Professor Brigitte Boothe, die mit der Erzählanalyse JAKOB ein Werkzeug geschaffen hat, das mir eine faszinierende Projektarbeit im Schnittpunkt von Narrativik, Linguistik und Informatik ermöglichte. Sie hat das Lexikonprojekt immer wohlwollend unterstützt und mir grosse Freiheit bei der Entwicklung meiner Ideen gelassen. Herr Professor Arnulf Deppermann gab mir mit seinem breiten sprachwissenschaftlichen Fachwissen wertvolle Hinweise und Anregungen für die Fertigstellung der Arbeit.

Tatkräftige Unterstützung erhielt ich von Simon Clematide vom Institut für Computerlinguistik sowie von Bernhard Distl und Jonas Huber, die das Lexikonprojekt ab 2007 technisch umgesetzt und betreut haben. Mein Dank geht auch an die Kolleginnen und Kollegen der Abteilung Klinische Psychologie, Psychotherapie und Psychoanalyse und an die Studierenden, die in Lehrveranstaltungen, Projektgruppen, Praktika und Lizentiatsarbeiten die Entwicklung des Lexikons und der Erzählanalyse unterstützt haben. Den Professoren Ulrich Heid und Patrick Hanks danke ich für lexikografische Beratung und Ideen.

Ein grosser Dank für die Unterstützung geht an meine Familie, die es mir ermöglicht hat, mich über längere Zeit intensiv mit meinem Projekt auseinander zu setzen.

Zürich, April 2011

Marc Luder

Zusammenfassung

Welche sprachlichen Konstruktionen und Gesprächspraktiken verwenden Klient und Therapeut, um sich im Therapiegespräch verständlich zu machen und um in der Interaktion Bedeutung/Sinn herzustellen? Wie werden Wörter und Ausdrücke im Gespräch mit Bedeutung versehen und wie verstehen Zuhörer das von Sprechern Gemeinte?

In der vorliegenden Arbeit wird ein elektronisches Konstruktionslexikon entworfen, mit dem Wörter und Ausdrücke mit linguistischen und extralinguistischen Informationen versehen abgebildet werden können. Im Hintergrund steht das erzählanalytische Verfahren Erzählanalyse JAKOB und die damit verbundene Analyse und Interpretation der lexikalischen Wahlen in den Erzählungen von Psychotherapiepatienten.

Konzepte der Konstruktionsgrammatik erweisen sich als geeignet für die Modellierung von Lexikoneinträgen, die aus in Mustern eingebetteten Einzelwörtern, vor allem aber aus variablen oder fixen Ausdrücken und Redewendungen bestehen. Eine Konstruktion ist eine Form-Bedeutungseinheit und umfasst Merkmale von Morphologie über Syntax und Semantik bis hin zu pragmatischen und funktionalen Eigenschaften der Wortverbindung, und des Diskurskontextes. Der Diskurskontext erweist sich dabei als der Knackpunkt und führt immer wieder zur Frage, welche (kognitiven) Prozesse für die Auswahl und Kombination von Wortverbindungen beim Meinen und Verstehen verantwortlich sind. Einerseits verfügen wir über ein grosses Repertoire von schnell abrufbaren vorgefertigten Phrasen, von konventionalisierten Wortkombinationen bis zu Metaphern und Idiomem, andererseits können wir diese Ausdrücke oder Teile davon im Sprachgebrauch auch auf kreative Weise neu kombinieren.

Für den Aufbau des Konstruktionslexikons werden Gesprächsausschnitte aus Transkripten mit gesprächsanalytischen Methoden untersucht; korpuslinguistische Verfahren belegen das Vorkommen der Wortverbindungen in den Gesprächskorpora in Bezug auf Häufigkeit, Kontext und weitere Eigenarten des Sprachgebrauchs. Aus diesen Befunden werden Hypothesen über die Funktionen einer Wortverbindung gebildet und die Lexikoneinträge modelliert. Datengrundlage sind drei Gesprächskorpora mit Psychotherapiegesprächen (ca. 500 transkribierte Stunden). Zu Vergleichszwecken werden ausserdem Daten aus dem Archiv für gesprochenes Deutsch, aus dem Schweizer Textkorpus und aus einem grossen deutschsprachigen Internet-Korpus verwendet.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Allgemeine Überlegungen	1
1.2	Forschungsinteressen und Fragestellung	6
1.3	Aufbau der Arbeit	10
2	Psychotherapie, Erzählen, Narrativ	15
2.1	Narrativ und Narrativik, Erzählanalyse	16
2.2	Das Narrativ in der Psychotherapie	22
2.3	Die Erzählanalyse JAKOB	24
2.3.1	Modellierungsfunktionen	25
2.3.2	Anwendung	26
2.3.3	Erzählung und Dramaturgie	27
2.3.4	Kodierung und Lexikon	30
2.4	Lexical Choice	33
2.5	Text und Verknüpfung	35
2.6	Fazit zu Narrativ und Erzählanalyse	36
2.7	Textanalyse in der Psychotherapie	37
2.7.1	Diktionäre	38
2.8	Fazit zu diktionsbasierten Verfahren	42
3	Linguistik und Gesprächsanalyse	43
3.1	Interaktionale Linguistik	45
3.2	Gesprächslinguistik, Gesprächsforschung	47
3.2.1	Konversationsanalyse, Conversation Analysis	50
3.2.2	Gesprochene-Sprache-Forschung	55
3.2.3	Beispiele interaktiver Gesprächspraktiken	57
3.3	Bedeutungskonstitution	59
3.3.1	Unterspezifizierung durch Implicitness	61
3.3.2	Unterspezifizierung durch Nicht-Determiniertheit	62
3.3.3	Unterspezifizierung durch Inkompatibilität	63
3.3.4	Bedeutungskonstitution und Konstruktionsgrammatik	63
3.3.5	Emergent Grammar	64

3.4	Satzsemantik - Pragmatik	65
3.5	Situation – Gesprächssituation	66
3.6	Fazit zum Kapitel Gesprächsforschung und Interaktionale Linguistik	68
4	Phraseologie	71
4.1	Einführung und Überblick	71
4.2	Mehrwortausdrücke, Multiword Expressions	75
4.2.1	Qualities of Multiword Expressions	80
4.2.2	Zum Begriff «Kollokation»	81
4.3	Idiome und Metaphern	83
4.3.1	Idiome	84
4.3.2	Metaphern	85
4.3.3	Metonymie	88
4.4	Phraseologie und Computerlinguistik	91
4.4.1	Word Sense Disambiguation (WSD)	94
4.5	Phraseologie und Psycholinguistik	95
4.6	Verben, verbale Phraseologismen	98
4.6.1	Klassifikationssysteme für Verben	98
4.6.2	Funktionsverben und Funktionsverbgefüge	103
4.7	Formulaic language, Formelhaftes	105
4.8	Fazit zum Kapitel Phraseologie	109
5	Korpuslinguistik	113
5.1	Statistische Assoziationsmasse in der Korpuslinguistik . . .	115
5.1.1	Mutual Information	116
5.1.2	logDice Score	116
5.2	Fazit zum Kapitel Korpuslinguistik	117
6	Konstruktionsgrammatiken	119
6.1	Funktionale Grammatiktheorien	120
6.2	Konstruktionsgrammatik	123
6.2.1	Was sind Konstruktionen	124
6.3	Hauptrichtungen der Konstruktionsgrammatik	126
6.4	Grundannahmen Konstruktionsgrammatik	130
6.5	Konstruktionsgrammatik und Interaktionale Linguistik . . .	131
6.6	Neuere Entwicklungen der Konstruktionsgrammatik	134
6.7	Von der Konstruktionsgrammatik zur Diskursgrammatik . .	135
6.8	Grammatikalisierung	140
6.9	Konstruktionsgrammatik: Darstellungsformalismen	140

6.10	Fazit zur Konstruktionsgrammatik	143
6.11	Corpus Pattern Analysis	145
6.11.1	Notation im Pattern Dictionary	149
6.11.2	Vergleich CPA – CxG – FrameNet	150
6.11.3	Fazit zur Corpus Pattern Analysis	151
6.12	Lexical Priming	152
6.12.1	Fazit zum Kapitel Lexical Priming	155
7	Lexikon und Lexikographie: Konzepte	157
7.1	Das JAKOB-Lexikon	158
7.1.1	Aufbau des bisherigen JAKOB-Lexikons	159
7.2	Bestehende Mehrwort-Lexika	160
7.2.1	WordNet	161
7.2.2	GermaNet	162
7.2.3	FrameNet	163
7.2.4	Das UWV-Analysemodell	165
7.2.5	Idiomatisierte verbale Phraseologismen (IVP)	168
7.2.6	Phraseo-Lex	169
7.2.7	Dutch Electronic Lexicon of Multiword Expressions	169
7.2.8	OLIF	171
7.2.9	Valenzwörterbuch deutscher Verben: VALBU	173
7.3	Ontologien, Texttypen, Diskursmuster	174
7.3.1	Makroebene: Domain, Genre, Subject Field	175
7.3.2	Meso- und Mikroebene: Gesprächstypen und lokale Muster	176
7.3.3	Semantische Ontologien	181
7.4	Diverse Konzepte zum Lexikon	183
7.4.1	Nominalität vs. Verbalität	183
7.4.2	Konnotation im Lexikon	185
7.5	Fazit zum Kapitel Lexikographie	187
8	Daten und Methoden	189
8.1	Gesprächsdaten, Korpora	190
8.1.1	Liste der verwendeten Korpora und praktisches Vorgehen	191
8.1.2	WIL: Korpus Wilma	191
8.1.3	AMA: Korpus Amalie	192
8.1.4	GUS: Korpus Gustav	193
8.1.5	AGD: Archiv für Gesprochenes Deutsch	193
8.1.6	CHTK: Schweizer Textkorpus	194

8.1.7	deWAC: Deutschsprachiges Korpus der Sketch Engine	195
8.1.8	Weitere Korpora	195
8.2	Lexika und MWE-Sammlungen	196
8.3	Methoden	197
8.3.1	Gesprächsanalyse	199
8.3.2	Korpusanalyse und Sketch Engine	202
8.3.3	Corpus Pattern Analysis	204
8.4	Fazit zum Kapitel Daten und Methoden	204
9	Durchführung - Lexikonerstellung	207
9.1	Neue Anforderungen an das JAKOB-Lexikon	208
9.1.1	Lexikonaufbau: Lesarten	209
9.1.2	OLIF Eigenschaften	210
9.1.3	Neue Merkmale im Lexikon	217
9.1.4	Weitere Überlegungen zum neuen Lexikon	228
9.1.5	Darstellung der Konstruktionen	230
9.2	Analysebeispiele, Konstruktionen	231
9.2.1	Schematische Konstruktion: <lesen>	232
9.2.2	Kommunikationskonstruktion: <an den Karren fahren>	234
9.2.3	Konstruktion <Sorge tragen>	244
9.2.4	Konstruktionen mit «Zeug»	249
9.2.5	Emotionskonstruktionen mit «verjagen», «explodieren» und «Dampf»	258
9.2.6	Emotionskonstruktionen mit «Angst»	266
9.3	Bemerkungen zu den Beispielkonstruktionen	272
9.3.1	JAKOB-Lexikon online	273
10	Schluss: Fazit und Ausblick	275
10.1	Korpora	275
10.2	Das Konzept der Konstruktion	277
10.3	Lexikonaufbau: Kritik und Vorschläge	277
10.3.1	Verknüpfungen (cross references)	278
10.3.2	Verbpatterns, Bauplan	279
10.3.3	Pragmatische Attribute	280
10.3.4	Diskurs über den Satz hinaus	280
10.4	Dynamische Bedeutungskonstitution	281
10.4.1	Meaning Construction	282
10.4.2	Norms and Exploitations	283
10.4.3	Das Paradoxon des Formelhaften	283
10.5	Erzählanalyse JAKOB – Computeranwendung	285

11 Literaturverzeichnis	287
A Anhang	307
A.1 Abkürzungsverzeichnis	307
A.2 Transkriptionsregeln	308
A.3 Notationsregeln im JAKOB-Lexikon	309
Index	313
Konstruktionsindex	315

1 Einleitung

1.1 Allgemeine Überlegungen

Welche sprachlichen Konstruktionen und Gesprächspraktiken verwenden Klient und Therapeut, um sich im Therapiegespräch verständlich zu machen und gegenseitig *Bedeutung/Sinn* zu konstruieren? Allgemeiner: Wie entsteht Bedeutung im Gespräch, wie versteht der Zuhörer das vom Sprecher Gemeinte? Der Ausgangspunkt der vorliegenden Arbeit ist die *Erzählanalyse JAKOB*. Das heisst, die Forschungsfragen und Probleme, die aufgeworfen werden, und die Lösungen, die angestrebt werden, sind aus der täglichen Arbeit mit der Erzählanalyse JAKOB entstanden. Als Erzählungen werden hier so genannte Alltags- oder Wirklichkeitserzählungen (Boothe, 2009; Klein & Martinez, 2009) betrachtet, die aus verschiedenen Kontexten stammen, vor allem aber aus psychotherapeutischen Gesprächssituationen.

(1) Beispielerzählung: «Kein Krankenhausbesuch bei Vater»¹

- 1 und dann äh kam dann der Freitagabend
- 2 und dann rief mich meine Mutter an
- 3 weil sie meinen Vater ins Krankenhaus bringen musste
und bezog mich gerade schon so ein so in das
äh machte so schon einen festen Besuchsrythmus
und verfügte wieder voll über meine Zeit
- 4 wann ich jetzt da äh meinen Vater besuchen müsse
im Krankenhaus
- 5 und äh dann geriet ich in einen grausamen Konflikt hinein
- 6 weil äh einerseits habe ich meinen Vater ja sehr gern
- 7 und andererseits wehrte ich mich mit jeder Faser dagegen
so mh mh jetzt so wieder so einbezogen wieder so verplant zu
werden, oder
- 8 und äh das machte mir dann überhaupt das ganze Wochenende
auch noch sehr Mühe
- 10 weil äh ich we ich ich wehrte äh mich dann bei meiner Mutter
und sagte

¹In der Erzählanalyse JAKOB werden die Erzählungen segmentiert (einfache Subjekt-Prädikatverbindungen), die Segmente nummeriert.

- 11 ich könne jetzt nicht , oder
12 ich hätte jetzt keine Zeit

Eine Frau erzählt in der Therapiesitzung eine Begebenheit aus der Vergangenheit. Als Erzählerin schildert sie eine Szene, bei der sie sich nicht verstanden fühlte und in der sie in einen Loyalitätskonflikt geriet. Die Erzählerin lässt einerseits die Ich-Figur auftreten und das Ereignis vorführen und nimmt andererseits auch als Erzählerin dazu Stellung (Boothe, 2011). Die Präsentation der Erzählung reinszeniert deshalb nicht nur eine vergangene Situation, sondern steht auch im Dienst der Gegenwart und dient der Erzählerin dazu, persönliches Erleben auszudrücken, sich in der therapeutischen Interaktion zu positionieren und vom Gegenüber eine Stellungnahme zu bekommen. Die Erzählerin führt Regie, erzeugt mit dramaturgischen Mitteln Spannung und führt den Zuhörer von einer Ausgangssituation zu einer Ergebnissituation. Erzählungen in der Psychotherapie ermöglichen einen Zugang zur erlebten Realität von Klienten oder Patienten und verweisen auf deren Konflikt- und Beziehungsdynamik (Arboleda et al., 2010).

Die Analyse von Erzählungen oder Narrativen kann von verschiedensten wissenschaftlichen Standpunkten und mit unterschiedlichen Forschungsparadigmen angegangen werden. Die Theorien und Methoden reichen von der konversationsanalytisch orientierten Untersuchung von Mikrostrukturen in kleinsten Gesprächseinheiten auf der einen Seite bis zur breit angelegten Analyse von grossen Textkorpora zum Zweck der empirischen Erforschung von rekurrenten Textmustern und Kollokationen auf der andern Seite.

Mit der Erzählanalyse JAKOB werden Gesprächsausschnitte untersucht, die eine episodische Struktur aufweisen. Erzähler stellen ihre Welt und ihre Wünsche und Ängste in *Episoden* szenisch dar. Die dabei verwendeten Mittel der sprachlichen Inszenierung können Aufschluss geben über innere Beweggründe und Konflikte. Die *lexikalische Wahl* ist ein Teil der sprachlichen Inszenierung und hat in der Erzählanalyse JAKOB eine besondere Funktion: Die Wortwahlen werden auf dem Hintergrund eines psychoanalytisch orientierten Kategoriensystems kodiert; die Zuordnung der Kategorien zu einzelnen Wörtern und Ausdrücken ist im so genannten *JAKOB-Lexikon* festgelegt, einem online zugänglichen elektronischen Wörterbuch. Im Idealfall kann also das von der Erzählerin verwendete Vokabular im Lexikon nachgeschlagen und mit den darin abgebildeten Kategorien kodiert werden. Dieses Vorgehen ist einerseits begrenzt durch den Lexikonumfang und andererseits durch Eigenschaften der Wortbedeutung, die situations- und kontextabhängig variiert und deshalb nicht eindeutig festgelegt werden kann. Einer der Arbeitsschritte in der Anwendung der Erzählanalyse

JAKOB besteht also in der *Kodierung* der verwendeten Wörter und Ausdrücke, speziell der *Verben*. Das Kategoriensystem für die Verben besteht aus 93 verschiedenen, psychodynamisch und psychoanalytisch motivierten Kategorien.

Die Tatsache, dass in der Erzählanalyse JAKOB ein Kodiersystem mit festgelegten Kategorien verwendet wird und dass diese Kategorien in einem Lexikon verankert sind, weist über die Erzählanalyse hinaus in das Feld der allgemeinen Textanalyse und zu linguistischen Fragestellungen (Boothe, 1989). Das Kodiersystem der Erzählanalyse JAKOB konzentriert sich hauptsächlich auf die *Verben*, die lexikalische Bedeutung fokussiert deshalb auf die *sprachlichen Handlungen*, die sich in den lexikalischen Setzungen zeigen. Dieser Ansatz geht in der Psychologie unter anderem auf Donald Spence zurück, der innerhalb der Psychotherapieforschung auf das Phänomen *lexical choice* schon zu Beginn der 1980er Jahre aufmerksam machte (Spence, 1980). Ausgehend von Freuds Gedanken zu Versprechern und weiteren sprachlichen Fehlleistungen (Freud, 1904), die vor allem das Vergessen und das falsche Erinnern zum Thema hatten, fragte sich Spence, ob nicht auch die lexikalische Wahl der Wörter bestimmten Zielen dienen könnte.

Mit der lexikalischen Wahl steuern Sprecher bewusst und unbewusst die Bedeutung, die sie mit ihren Äußerungen transportieren wollen. Damit stellt sich die Frage, wie *Bedeutung* entsteht, wie sie produziert und rezipiert wird, ob sie an einzelnen Wörtern festgemacht werden kann, oder ob die Herstellung von Bedeutung eher auf Ausdrucks-, Satz- und Diskursebene stattfindet. «I don't believe in word senses» sagt Kilgarrieff (2008) und meint damit, dass das Zuweisen von Bedeutung an einzelne lexikalische Items nicht sinnvoll und nützlich ist. In der *Lexikographie* hat diese Erkenntnis seit längerem dazu geführt, dass Wörterbücher entwickelt werden, die das lexikalische Item in den Kontext der usuell mit ihm verwendeten Wörter und Ausdrücke stellen, was vor allem im Zweitsprach-Erwerb nützlich ist. Deppermann (2006e) führt den Begriff der *Bedeutungskonstitution* ein, um das Entstehen von Bedeutung im Gespräch zu beschreiben, und er betont aus gesprächsanalytischer Sicht, dass Bedeutung nur im Kontext entstehen kann. Das traditionelle «Einwort-Lexikon» wird also von verschiedener Seite in Frage gestellt, wenn es darum geht, Wörtern Bedeutung zuzuschreiben.

Im Gegensatz zu diesen theoretischen Überlegungen stehen so genannte *Diktionäre*, spezialisierte psychologische Wörterbücher, die den Wortschatz einer bestimmten Wissensdomäne oder bestimmter psychologischer Themenbereiche abdecken und ohne Berücksichtigung des Kontextes für Textanalysen verwendet werden. Bekannt sind etwa die Gottschalk-Gleser Angstskalen oder das Dresdner Angstwörterbuch, beides Instrumente, die

mit Computerhilfe auf Grund des Vorkommens von bestimmten, im Diktionär festgelegten *Wörtern* in Gesprächen auf den Ausdruck von Angst schliessen. Auch das Analyseinstrument *Linguistic Inquiry and Word Count* (LIWC) des amerikanischen Psychologen James Pennebaker (Pennebaker & Francis, 2001) verwendet ein Kategoriensystem, das sich am Vorkommen *einzelner Wörter* in Texten orientiert und daraus Variablen berechnet. In der Psychotherapieforschung bekannt und (validiert) ist das Emotions- und Abstraktionsvokabular, das aus dem Wortschatz aus Therapiegesprächen Hinweise auf Phasen des emotionalen Erlebens oder der Abstraktion/Reflektion hinweist (z.B. Mergenthaler, 1997).

Die zweite Frage, die sich gegenüber einer lexikalischen Auflistung von Wort- oder Ausdrucksbedeutungen stellt, ist grundsätzlicher Natur: Können Wort- und Ausdrucksbedeutungen a priori in einem Lexikon festgeschrieben werden? In der *interaktionalen Linguistik* und in der *Gesprächsanalyse* wird argumentiert, dass die Bedeutung von Äusserungen interaktiv entsteht und ausgehandelt wird und deshalb (innerhalb bestimmter Grenzen) nicht zum voraus bestimmbar ist²; Wortbedeutungen können deshalb nicht kontextfrei (unabhängig von der Situation) und isoliert von der Interaktion festgelegt werden. Die Gesprächsstruktur wird gemeinsam von den Gesprächsteilnehmern erarbeitet. Auch die *Konversationsanalyse* (die hier unter den Begriff der *Gesprächsanalyse* subsumiert wird) betont die Wichtigkeit von Interaktion und Kontext für den Analyseprozess und verlangt grundsätzlich ein strikt induktives Forschungsvorgehen ohne a priori vorgegebene Kategorien (Hakulinen & Selting, 2005a). Allerdings räumen die Autoren ein, dass z.B. syntaktische Strukturen und lexikalische Items und Kategorien kaum als nicht vorgegeben vorgestellt werden können. Sie sind aber im Interpretationsprozess zu hinterfragen. Die *Interaktionale Linguistik* untersucht, welche linguistischen Ressourcen verwendet werden, um bestimmte Gesprächsabläufe zu erzeugen und interaktionale Funktionen zu erfüllen, und umgekehrt, welche Gesprächsstrukturen und interaktionalen Funktionen durch die Verwendung bestimmter linguistischer Formen gefördert werden (Couper-Kuhlen & Selting, 2001).³

Das Phänomen der *Emergenz* wird von Hopper (1987) und anderen Autoren (Auer, 2007; Günthner, 2007b; Selting & Couper-Kuhlen, 2000) als

²Völlig willkürlich entsteht Bedeutung allerdings kaum, sonst wäre eine Verständigung nicht möglich.

³i. what linguistic resources are used to articulate particular conversational structures and fulfil interactional functions?

ii. what interactional function or conversational structure is furthered by particular linguistic forms and ways of using them? (Couper-Kuhlen & Selting, 2001)

wichtige Eigenschaft des Gesprächsverlaufs bezeichnet.⁴ Emergenz heisst in dieser Sichtweise, dass im Diskurs die Bedeutung der einzelnen Äusserungen laufend (emergent) hergestellt wird. Gesprächsanalyse und interaktionale Linguistik untersuchen vor allem *gesprochene Sprache*, und es wird Wert darauf gelegt, dass empirische Daten als Grundlage für die Analysen verwendet werden (im Gegensatz zu den artifiziellen Beispielsätzen der klassischen Linguistik). Die Disziplin, die sich ausschliesslich um Phänomene der gesprochenen Sprache kümmert, heisst denn auch *Gesprochene Sprache-Forschung* (Schwitalla, 2001, 2003).

Reale Text- und Gesprächsdaten sind in *Korpora* repräsentiert. Die Theorien und Methoden der *Korpuslinguistik* können dazu beitragen, die in kleinen Gesprächsausschnitten mit gesprächsanalytischen Methoden gefundenen Erkenntnisse auf grössere Textmengen zu beziehen und Besonderheiten des Sprachgebrauchs auch empirisch zu belegen (oder überhaupt zu finden). Der englische Korpuslinguist John Sinclair leistete Pionierarbeit für eine funktionale Sicht auf sprachliche Phänomene, indem er auf Grund von Korpusuntersuchungen die *Musterhaftigkeit* von Sprache postulierte. Bedeutung entsteht zu einem grossen Teil durch und in phraseologischen Textmustern (Moon, 2008).

Die Forschungsergebnisse zur *Phraseologie* zeigen eine ähnliche Richtung auf wie oben bei der Gesprächsanalyse und Interaktionalen Linguistik angesprochen: Wörter und Ausdrücke können nur im Kontext von ganzen Äusserungen eine spezifische Bedeutung bekommen. Die Phraseologie beschäftigt sich mit *Phraseologismen* oder *Mehrwortausdrücken*,⁵ die von einfachen Wortkombinationen bis zu ganzen Sätzen reichen und umfasst sowohl Aspekte der Morphologie und Syntax als auch der Semantik und der pragmatischen Gebrauchsbedeutung.

In den Theorien der *Konstruktionsgrammatik* (construction grammar) sind die grundlegenden Einheiten der Sprache *Verbindungen aus Form und Bedeutung* (Form-Bedeutungspaare) und bestehen aus Ausdrücken oder Phraseologismen, die ganz oder teilweise lexikalisiert sein können. Die *Konstruktionen* stehen in unmittelbarem Bezug zu ihren Funktionen im Sprachgebrauch. Das sprachliche Wissen ist auf der Ebene von Konstruktionen repräsentiert und besteht aus Ausdrücken verschiedenster Komplexität. Die Theorien der Konstruktionsgrammatik sind in den 1980er Jahren im Umfeld der kognitiven Linguistik entstanden (Lakoff, 1987).

⁴Vgl. die *Emergent Grammar* (Hopper, 1987).

⁵Englisch: *multiword expressions* (MWE) oder *multiword units* (MWU).

Korpusanalytische Verfahren wie die *Corpus Pattern Analysis* versuchen, die Bedeutung von Wörtern und Ausdrücken auf Grund der Muster ihres Vorkommens in grossen Textkorpora zu erschliessen (Hanks, 2008a). Die Wortbedeutung ergibt sich aus der Verwendung der Wörter in bestimmten syntagmatischen Mustern und Satzkontexten, d.h. jedes Wort ist mit einem oder mehreren syntagmatischen Mustern⁶ assoziiert.

Angestrebt wird im *theoretischen Teil* der vorliegenden Arbeit eine Zusammenfassung der wichtigsten Ideen der genannten Einzeldisziplinen und der bestehenden Zusammenhänge. Wie die Übersicht zeigt, werden zum Teil dieselben oder ähnliche Phänomene untersucht, und oft werden dabei die Erkenntnisse aus verwandten Disziplinen zu wenig beachtet. Ein besonderes Anliegen der vorliegenden Arbeit ist deshalb auch die Integration der verschiedenen Ansätze im Lexikonprojekt.

Im zweiten, praktischen Teil geht es darum, das bestehende *JAKOB-Lexikon* zu einer reichhaltigen elektronischen Lexikonressource auszubauen. Lexikoneinträge sollen mit so viel Zusatzinformationen ausgestattet werden, dass sie eine bessere Zuordnung der im Gespräch verwendeten Wörter und Ausdrücke zu lexikalischen Kategorien (und damit auch zu einer spezifischen Kodierung in der Erzählanalyse JAKOB) ermöglichen. Im Hintergrund des Projekts steht die *Erzählanalyse JAKOB* und die damit verbundene Analyse und Interpretation der lexikalischen Wahlen in den Erzählungen von Psychotherapiepatienten. Ziel des Projektes ist ein Prototyp für ein elektronisches Konstruktionslexikon, das sowohl vom menschlichen Lexikonbenutzer als auch von NLP-Applikationen⁷ durchsucht werden kann.

1.2 Forschungsinteressen und Fragestellung

Die Forschungsinteressen für das vorliegende Projekt können in der folgenden anspruchsvollen Formulierung von Hakulinen und Selting (2005a) zusammengefasst werden:

the task of the researcher is to deconstruct the activity that is interpreted in a holistic way by recipients into constitutive cues, structures and devices, and to reconstruct the way in which it is thus systematically made interpretable for recipients. (Hakulinen & Selting, 2005a, S. 3)

⁶In der CPA handelt es sich ausschliesslich um Verbpatterns.

⁷NLP: natural language processing.

[Übersetzung M.L.] Die Aufgabe des Forschers ist es, die [sprachlichen] Aktivitäten, die von den Rezipienten ganzheitlich interpretiert werden, in konstitutive Besonderheiten, Strukturen und Einheiten zu zerlegen, und darauf aufbauend zu rekonstruieren, wie diese [Bausteine] systematisch dazu beitragen, dem Rezipienten diese Aktivitäten verständlich zu machen.

Wie können die Erkenntnisse aus Psychologie, Gesprächsforschung, Phrasologie, Korpuslinguistik, Konstruktionsgrammatik und Lexikographie in einem *Konstruktionslexikon* (Konstruktikon) umgesetzt und für die Erzählanalyse JAKOB fruchtbar gemacht werden? Welche Ziele sollten für die weitere Entwicklung von Lexikon und Erzählanalyse verfolgt werden? Die neuen Lexikonkonzepte werden exemplarisch anhand von mehreren Konstruktionen vorgestellt; eine grossflächige Implementierung ist im vorliegenden Rahmen nicht möglich. Die technisch wichtigen Neuerungen für das Lexikon wurden bereits ab dem Jahr 2008 vorgenommen (neue Datenbankstruktur, OLIF-Schema etc., siehe Kap 9.1.1), so dass das Interesse sich in der vorliegenden Arbeit auf inhaltliche und konzeptuelle Fragen der Umsetzung konzentrieren kann.

Wie funktioniert die Erzählanalyse und welches ist die Rolle des Lexikons? Die folgende kurze Beschreibung fasst die wesentlichen Schritte zusammen; ausführlicher wird die Erzählanalyse JAKOB im Kapitel (2.3) vorgestellt, das Lexikon in den Kapiteln (7.1) und (9).

Erzählungen werden nach bestimmten Kriterien aus Transkripten von Therapiegesprächen extrahiert (episodische Darstellung und sequentieller Ablauf von mindestens zwei Handlungen mit markiertem Anfang und Ende). Diese Erzählungen werden anschliessend segmentiert, d.h. in Subjekt-Prädikat-Einheiten oder einfache Sätze zerlegt und so in die Erfassungsmaske der internetbasierten JAKOB-Anwendung eingegeben. Die folgenden Schritte der textnahen Analyse der *Erzähldynamik* umfassen die Markierung von Segmenten, die *episodisch* sind, d.h. die der unmittelbaren Darstellung der Szene dienen, gegenüber solchen, die Rahmenbedingungen und Situation beschreiben und deshalb als *nicht episodisch* bezeichnet werden. Ausserdem werden drei Phasen im Ablauf der Erzählung markiert: die Einleitung, die die *Startdynamik* konstituiert, eine zweite Phase, die die *Entwicklung* der erzählten Handlung beschreibt und schliesslich die *Ergebnisphase*. Im nächsten Schritt erfolgt die bereits beschriebene *Kodierung* des verwendeten Vokabulars mit Computerhilfe auf Grundlage des Lexikons. Falls die zu kodierenden Wörter und Ausdrücke nicht im Lexikon vorkommen, muss die Kodierung manuell durchgeführt werden, d.h. diesen Aus-

drücken müssen aus der *Kategorienliste* geeignete Kategorien zugeordnet werden.

Die gefundenen Kodierungen dienen dazu, die weiteren Schritte der Erzählanalyse zu erleichtern und *Interpretationshilfen* zu geben. So dienen Codes beispielsweise dazu, die *Selbstpräsentation* der Erzählerin gegenüber ihrer Präsentation der Ich-Figur zu charakterisieren und die *soziale Involviertheit* dieser Rollen durch die mit Hilfe der Verbcodes erschlossenen Ausprägungen auf den Achsen *Macht*, *Nähe* und *Autonomie* zu bestimmen. Anschliessend helfen die Verbcodes mit, die *Startdynamik* als Grundlage für die folgende hypothetische Formulierung einer so genannten *Spielregel* für den weiteren Verlauf der Erzählung zu finden. Diese wiederum dient als Ausgangspunkt für die Formulierung einer *Konflikthypothese*, die sich auf die *Wünsche*, *Ängste* und *Abwehrmechanismen* bezieht, die aus der Erzählung erschlossen werden können.

Das vorliegende Lexikonprojekt beschäftigt sich mit einem Teilschritt der Erzählanalyse JAKOB, nämlich mit der *Kodierung* des Erzählvokabulars, besonders der Verben. Die Kodierung erfolgt auf Grund einer umfangreichen, a priori bestehenden *Kategorienliste*. Das *JAKOB-Lexikon* dagegen ist eine Wortliste, deren Einträge bereits mit den Kategorien aus der erwähnten Liste kodiert sind. Theoretisch ist also jedem Wort im Lexikon ein Code oder eine Codekombination zugeordnet; da das Lexikon aber aktuell nur wenige Tausend Einträge beinhaltet, gibt es viele Lücken. Ein weiteres mit diesem vereinfachten Blick auf das Lexikon zusammenhängendes Problem ist die Polysemie. Ein Wort kann mehrere *Bedeutungen* haben, und es muss deshalb klar sein, welche Bedeutungsvariante ein Lexikoneintrag repräsentiert. Die zwei wichtigsten Konsequenzen, die aus den erwähnten Beschränkungen resultieren, sind erstens eine Erweiterung des Lexikoneintrags vom Einzelwort zum *Ausdruck*, der ein oder mehrere Wörter umfassen kann, zweitens ist eine differenzierende Beschreibung polysemer Lexikonausdrücke in der Art erforderlich, dass eine Zuordnung von Ausdrücken aus den Erzählungen zu einem Lexikoneintrag mit der passenden Bedeutung möglich ist.

Die Vorteile des Kodierens auf der Basis des Lexikons liegen darin, dass die Zuordnung der Codes zum Erzählvokabular reliabler durchgeführt werden kann und dass dadurch die Kodierung bis zu einem gewissen Grade objektiviert wird. Für die nachfolgenden interpretativen Schritte der Erzählanalyse ist damit eine vergleichbare Ausgangslage geschaffen. Wichtig ist schliesslich auch der ökonomische Gesichtspunkt: die automatische Kodierung spart bei der Analyse viel Zeit gegenüber dem manuellen Vorgehen. Dies wird allerdings etwas relativiert, weil die automatisch eingesetzten

Codes überprüft werden müssen, was ebenfalls zeitaufwendig ist und Sachkenntnis erfordert. Auch dies ist ein Argument für die weitere Entwicklung und Verbesserung des Lexikons.⁸

Für die vorliegende Studie wird der Fokus von der Erzählung auf die ganzen *Gespräche* erweitert, um mehr Gesprächsmaterial zur Verfügung zu haben. Ziel ist es, in den vorgestellten Korpora aus Therapiegesprächen prototypische Beispiele von *Ausdrücken* (Konstruktionen)⁹ zu finden, die gesprächs- und kontextspezifische Bedeutung transportieren, diese als Lexikoneinträge mit einem reichhaltigen Set von lexikalisch nützlichen Eigenschaften zu versehen und die praktische Verwendbarkeit eines solchen Lexikons für das Verstehen (die Bedeutungsrezeption) von Gesprächsausschnitten zu prüfen. Dabei ist ein gewisser Kontext durch die Transkripte vorgegeben (Genre Psychotherapiegespräch). Die oben dargestellte JAKOB-Kodierung ist ein einzelnes Merkmal im Lexikoneintrag, die Bedeutung des Items wird jedoch nicht ausschliesslich am Code festgemacht, sondern an der Summe aller Eigenschaften. Die Kategorisierung ist also wesentlich feinkörniger als dies für die Kodierung in der Erzählanalyse nötig wäre. Das Forschungsinteresse beschränkt sich nicht nur auf die reliablere Kodierung, sondern lässt sich weiter fassen als Exploration der grundsätzlichen *Unterscheidungsmöglichkeiten von lexikalischen Items* auf Grund der Konstruktionsmerkmale. Lexikoneinträge gehen deshalb über die Ebene der Bedeutungszuweisung durch die JAKOB-Kodierung hinaus, d.h. für differenziertere Unterscheidungen stehen eventuell keine spezifischen Codes zur Verfügung. Das Ziel besteht trotzdem darin, Bedeutungsnuancen zu erfassen, auch wenn dafür auf der Codeebene nicht adäquate Unterscheidungen gemacht werden können.

Folgende *Forschungsfragen* sollen beantwortet werden:

1. Können *Konstruktionen* (Lexikoneinträge) mit den vorgeschlagenen Methoden aus den Korpusbeispielen erschlossen und adäquat dargestellt werden, und welches sind die Anforderungen dafür?
2. Können *gleich lautende oder ähnliche Konstruktionen* mit unterschiedlicher Bedeutung (Homonymie und Polysemie) im Lexikon so differenziert dargestellt werden, dass eine Unterscheidung auf Grund der Lexikonmerkmale möglich ist?

⁸Die automatische Kodierprozedur muss natürlich ebenfalls verbessert werden; dies gehört jedoch nicht zum vorliegenden Projekt.

⁹Lexikonitems werden in der vorliegenden Betrachtungsweise als *Konstruktionen* im Sinne der *Konstruktionsgrammatik* bezeichnet, siehe Kapitel (6).

3. Repräsentieren Lexikoneinträge eine *Grundbedeutung* (Kernsemantik), die als Grundlage für die *Bedeutungskonstitution* in der Gesprächssituation verwendet werden kann?
4. Wieviel *Kontext* müssen Lexikoneinträge enthalten? Wie dekontextualisiert dürfen sie sein?
5. Ist der Begriff der *Konstruktion* ein passendes Konzept für die Abbildung lexikalischer Phänomene?
6. Wie funktioniert die *Bedeutungskonstitution dynamisch* in der Interaktion? Können über diesen Prozess Erkenntnisse gefunden und als Einsichten oder Hypothesen für die Gesprächs- und Erzählanalyse gewonnen werden?
7. Kann die *automatische Kodierung* in der Erzählanalyse JAKOB mit den geplanten Lexikonergänzungen verbessert werden?

Die Projektziele und Forschungsinteressen bestehen also einerseits aus dem praktischen Anliegen, ein funktionierendes Konstruktionslexikon (oder mindestens einen funktionierenden Prototypen) zu erstellen, aber andererseits auch aus dem Interesse an den grundlegenden Funktionsweisen der Bedeutungsherstellung im Gespräch.

1.3 Aufbau der Arbeit

Die im Theorieteil dargestellten Verfahren und Konzepte reichen von der Erzählanalyse über die Gesprächsanalyse und die Interaktionale Linguistik bis zu Phraseologie, Korpuslinguistik, Konstruktionsgrammatik und Lexikographie. Der Ausgangspunkt ist die *Erzählanalyse JAKOB* und die Anwendung des *JAKOB-Lexikons* bei der Durchführung der Erzählanalyse. Kapitel (2) geht von den beiden Begriffen *Psychotherapie* und *Erzählanalyse* aus, um Theorien zu den Funktionen des Narrativs im Alltag und in Therapiegesprächen zu skizzieren und Anwendungen vorzustellen. Das *Narrativ* hat sich zu einer wichtigen Grundlage vieler interdisziplinär verknüpfter Forschungsansätze entwickelt. Narrative oder Erzählungen gelten als grundlegende menschliche Strategie, mit der sich das Individuum Sinn verschafft und sich in der Welt bewegt (Herman, Jahn & Ryan, 2005). *Sprachanalytische Verfahren* in der Psychotherapieforschung untersuchen Gespräche, Gesprächsausschnitte oder kürzere, als «Narrativ», «Erzählung» oder «small story» bezeichnete Gesprächs- und Textausschnitte und beruhen meistens

auf Kategoriensystemen; Texte und Gespräche werden nach verschiedenen Gesichtspunkten kategorisiert und kodiert. Linguistische und interaktionelle Phänomene werden anhand von sprachlichen Merkmalen dargestellt und erklärt. Oft werden dafür *Diktionäre* verwendet, Sammlungen von Wörtern, die als Marker für bestimmte (psychologische) Kennwerte dienen. Das Kapitel zum Narrativ enthält ausserdem eine kurze Einführung zur *Erzählanalyse JAKOB*, die den Hintergrund zum vorliegenden Lexikon-Projekt bildet.

Das Kapitel (3) stellt Forschungsrichtungen aus dem Umfeld der *Gesprächsforschung* und der *Interaktionalen Linguistik* vor. Die *Konversationsanalyse* entstand im Kontext von Soziologie und Ethnomethodologie in den 1960er Jahren in den USA und verfolgt das Ziel, aus der minutiösen Aufzeichnung und Untersuchung von Gesprächen Regeln und Verfahren der Gesprächsorganisation zu erschliessen (Brinker & Sager, 2006; Deppermann, 2006c). Der im deutschen Sprachraum gebräuchliche Begriff *Gesprächsanalyse* bezieht sich einerseits auf die aus dem amerikanischen Sprachraum stammende Konversationsanalyse, andererseits auf die seit den 1960er Jahren aufkommende *Gesprochene-Sprache-Forschung*. Ein dritter Forschungszweig, der die Gesprächsanalyse beeinflusst, ist die *Linguistische Pragmatik* (Brinker & Sager, 2006, S. 15). Im Unterschied zu den stark am Gesprächstext ausgerichteten narrativen Analyseverfahren interessiert sich die Gesprächsanalyse mehr für den Gesprächskontext und die Interaktion, d.h. für die situationalen und personellen Umstände, unter denen ein Gespräch abläuft. Ein wichtiges Thema in diesem Kapitel ist ausserdem die *Bedeutungskonstitution*: mit welchem sprachlichen Repertoire und mit welchen Techniken wird Bedeutung im Gespräch produziert und rezipiert?

Das Kapitel (4) beschäftigt sich mit *Phraseologismen* oder *Wortverbindungen* (Burger, 2007). Mehr oder weniger feste Wortverbindungen bestimmen einen grossen, wenn nicht den grössten Teil unserer Sprachproduktion und -rezeption, das *Meinen* und *Verstehen* (Steyer, 2004b). Die Phraseologie stützt sich für ihre Aussagen und Konzepte auf Wortverbindungen, wie sie im realen Sprachgebrauch vorkommen, nicht auf Intuition und Introspektion. Auch dieser Gesichtspunkt ist für die vorliegende Untersuchung wichtig, geht es doch darum, eine konkrete Sprache (Deutsch) im alltäglichen Gebrauch zu untersuchen, nicht eine virtuelle Universalsprache. Phraseologische Phänomene wie Idiome, Metaphern und Metonymie kommen in diesem Kapitel ebenso zur Sprache wie die Beziehungen der Phraseologie zu Computerlinguistik und Psycholinguistik. Eine wichtige Gruppe der Wortverbindungen bilden die verbalen Phraseologismen und die Funktionsverbgefüge. Ausserdem werden in kurzen Zusammenfassungen Grundzüge

der *Korpuslinguistik* und die Forschung zur *formulaic language* vorgestellt.

Im Kapitel (6) werden Konzepte der *Konstruktionsgrammatik* vorgestellt. Die Konstruktionsgrammatik gehört zu den funktionalen Grammatiken; ihre Grundgedanken und Hauptrichtungen werden mit Blick auf die Verwendung im Lexikonprojekt beschrieben. Ideen und Methoden der Konstruktionsgrammatik werden auch kombiniert mit Ideen und Verfahren der interaktionalen Linguistik, der Gesprächsanalyse, Phraseologie und Korpuslinguistik. Die Bezeichnung *Grammatik* darf hier nicht zu eng gesehen werden; Konstruktionsgrammatiken sind linguistische Theorien, die nicht nur grammatische Phänomene umfassen, sondern ebenso semantische, pragmatische und diskursive. Konstruktionen gelten als umfassende Beschreibungsrahmen für sprachliches Wissen (Deppermann, 2006a). Im Zusammenhang mit der Konstruktionsgrammatik wird auch das Phänomen der *Grammatikalisierung* erwähnt, und schliesslich werden *Darstellungsformalismen* für Konstruktionen vorgestellt.

Die *Corpus Pattern Analysis* und das dazu gehörige *Pattern Dictionary of English Verbs* (Hanks, 2008a) werden im Kapitel (6.11) vorgestellt und an Beispielen erläutert. Schliesslich wird die Theorie des *Lexical Priming* eingeführt (Hoey, 2005). Gemäss dieser Theorie entsteht die Bedeutung von Wörtern und Ausdrücken im Gebrauch und abhängig von der Häufigkeit des Gebrauchs in bestimmten/passenden Situationen. Die Bedeutung eines Wortes entwickelt sich durch seine Rezeptionsgeschichte (geschrieben und gesprochen). Das Wort wird durch Kontext und Kotext semantisch aufgeladen. Dasselbe gilt auch über das Wort hinaus für Ausdrücke und Sätze.

Wenn es darum geht, die Erkenntnisse aus Gesprächs- und Erzählanalyse, Phraseologie und Korpuslinguistik zu verbinden, kommt das Hauptthema des vorliegenden Projekts in Sicht, nämlich die Frage, wie ein elektronisches *Lexikon* aussehen könnte, das Ausdrücke aus dem allgemeinen oder aus einem speziellen Sprachgebrauch als Konstruktionen abbildet, das feste und variable Wortverbindungen (Phraseologismen) darstellen kann, und das diese Wortverbindungen auf Grund von Sprachkorpora charakterisiert. Nach einer Beschreibung des bisherigen *JAKOB-Lexikons* als Startpunkt werden im Kapitel (7) einige existierende elektronisch verfügbare *Mehrwort-Lexika* vorgestellt sowie *Typologien* und *Ontologien* für die Erfassung semantischer und pragmatischer Merkmale. Neben der Darstellung der theoretischen Konzepte geht es hier vor allem um die praktischen lexikographischen Aufgaben. Wie soll das *JAKOB-Lexikon* gestaltet werden, um den theoretischen Erörterungen und Anforderungen gerecht zu werden?

Das Kapitel (8) stellt die für die Analysen verwendeten *Korpora* vor. Dabei handelt es sich hauptsächlich um Transkripte aus Psychotherapiege-

sprächen, also aus dem Datenmaterial, für das die Erzählanalyse JAKOB ursprünglich entwickelt wurde. Zusätzliche öffentlich verfügbare Korpora werden als Referenz in die Untersuchung eingeschlossen, die sowohl Gesprächsdaten als auch Textdaten aus schriftlichen Quellen (Zeitungen, Internet) enthalten. Anschliessend werden die Methoden, die für die Aufbereitung des Lexikons verwendet werden, vorgestellt.

Das Kapitel *Durchführung* (9) zeigt an einer Anzahl prototypischer Beispiele die Fragen und Probleme auf, die sich beim Bau des Konstruktionslexikons ergeben, und entwickelt Lösungsvorschläge. Die Überlegungen zu den Konstruktionen, die Beispielanalysen und die Implementierungsvorschläge sollen dem praktischen Ziel der Umsetzung im JAKOB-Lexikon dienen. Das Kapitel (10) schliesst die Arbeit ab mit Kommentaren zu den gelungenen und weniger gelungenen Umsetzungen, zu gelösten und weiterhin ungelösten Problemen. Schliesslich folgen einige Empfehlungen für die weitere Arbeit mit der Erzählanalyse JAKOB und für weitere Projekte.

2 Psychotherapie, Erzählen, Narrativ

Psychotherapie und Erzählanalyse sind die beiden projektbezogenen Ausgangspunkte für eine Darstellung der Funktionen des *Narrativs* und der damit verbundenen Theorien und Anwendungen. Die Analyse von Patientengesprächen im Rahmen der Psychotherapieprozessforschung und der Anwendung des erzählanalytischen Verfahrens JAKOB bilden die Grundlage für die Betrachtung von Narrativen. Im Zentrum dieses Kapitels steht die Beschreibung von Verfahren, mit denen Texte oder Gespräche auf dem Hintergrund der Begriffe *Narrativ* und *Erzählung* als (monologische) sprachliche Produktionen analysiert werden, aber auch Verfahren, die therapeutische Prozesse aus allgemeinen Gesprächsdaten erfassen und erklären. Es soll eine Verortung der Begriffe *Narrativ* und *Erzählung* vorgenommen werden, die oft in ganz unterschiedlichen Bedeutungen verwendet werden, von der Bezeichnung für kurze episodische Erzählungen bis hin zur narrativen Darstellung und zum narrativen Verständnis ganzer Lebensabschnitte. Nach einer Darstellung von Studien und Analyseverfahren zum *Narrativ* speziell in der Psychotherapie wird die *Erzählanalyse JAKOB* und ihre Anwendung ausführlich vorgestellt und damit verbunden die spezielle Ausprägung des *Narrativs* in dieser Methode. Es folgt ein Exkurs zum Begriff der *lexikalischen Wahl* (lexical choice), der in der Erzählanalyse JAKOB eine besondere Rolle spielt und als Konzept für die Lexikonentwicklung wichtig ist. Schliesslich werden einige Textanalysewerkzeuge vorgestellt, die in der Psychologie und Psychotherapieforschung verwendet werden und auf so genannten *Diktionären*, einfachen spezialisierten Wörterbüchern, aufgebaut sind. Für das *Lexikonprojekt* wird damit die Ausgangslage skizziert, nämlich die Entstehung aus der praktischen Arbeit mit der Erzählanalyse JAKOB und mit einem Lexikon für die Kodierung der Verben in Form einer einfachen Verbliste.

Ausgehend von den strukturalistischen narrativen Theorien hat sich das *Narrativ* (in verschiedenen Ausprägungen des Begriffs) zu einer wichtigen Grundlage vieler interdisziplinär verknüpfter Forschungsansätze entwickelt. Narrative oder Stories sind der Ausdruck verschiedenster kognitiver und

kommunikativer Aktivitäten und gelten als grundlegende menschliche Strategie, mit der sich das Individuum Sinn verschafft und sich in der Welt bewegt (Herman et al., 2005). Die folgenden Schlagworte zeigen beispielhaft das breite Spektrum der mit dem Begriff Narrativ verbundenen Phänomene und die Relevanz für verschiedene Forschungskontexte (Herman et al., 2005)¹: *cognitive narratology, communication in narrative, biological foundations of narrative, artificial intelligence and narrative*. Die narrative Intelligenz bezieht sich auf das *episodische Wissen*, das in Erzählungen gespeichert und weiterentwickelt wird, der *narrative turn* bezeichnet eine Entwicklung in der Philosophie, Psychologie und weiteren Sozialwissenschaften, in Neurowissenschaft und Linguistik hin zu einer Sicht, in der das Narrativ als Ausdruck kognitiver und kommunikativer Aktivitäten eine wichtige Rolle spielt (Herman, 2003, S. 4–5).

2.1 Narrativ und Narrativik, Erzählanalyse

McLeod (1997) definiert *Narrativ* als eine Gedanken- und Diskurskategorie, die von andern Kognitions- und Kommunikationsformen grundsätzlich zu unterscheiden ist. Er braucht den Begriff Narrativ für den therapeutischen Diskurs als Ganzes, den Begriff Story für Darstellung einer spezifischen Episode im Sinne einer Alltagserzählung.

Eine prototypische Definition der Erzählung vom Typ *personal-experience narrative* stammt von Labov und Waletzky (1967). Demgemäss besteht die Struktur einer Erzählung aus Abstract, Orientierung, Komplikation (complicating action), Evaluation, Resultat und Coda. Wichtig ist die Feststellung, dass Narrative darüber hinaus Merkmale auf der Ausdrucks-, Satz-, und Diskursebene aufweisen (clause, sentence, discourse) (Herman, 2003, S. 8).

Elisabeth Gülich unterscheidet für das mündliche Erzählen die *narrative* und die *szenische Rekonstruktion* (Gülich, 2007a). Erzählen wird als *narrative Rekonstruktion* betitelt, während die «szenische Darstellung» als *nicht-narrative Form der Rekonstruktion* bezeichnet wird (S.36–37). Gülich kontrastiert die Erzählforschung mit der Konversationsanalyse: Die *Erzählforschung* fokussiert ihre Interessen auf den Erzähler und auf die Erzählung; in der Konversationsanalyse steht dagegen die koordinierte Aktivität aller Beteiligten im Vordergrund; die interaktiven Aspekte sind ebenso wichtig wie Text und Sprecher allein (S. 38), vgl. auch Georgakopoulou (2007).

¹Diese Begriffe repräsentieren das Programm der *Encyclopedia of Narrative Theory* (Herman et al., 2005).

Herman (2003) entwirft eine *interdisziplinäre narrative Theorie* (interdisciplinary narrative theory) und stellt diese in den grösseren Zusammenhang der Kognitionswissenschaften. Narrativ und Story und ihre wissenschaftliche Erforschung werden in diversen Kontexten verortet und beschrieben, so werden z. B. konstruierte Stories für die Durchführung von Experimenten verwendet, in der Psychotherapie werden Erzählungen analysiert, um Erkenntnisse über die Probleme der Patienten und über den therapeutischen Prozess zu gewinnen. Narrative werden mit einer Vielzahl von quantitativen und qualitativen Methoden untersucht und veranlassen Herman zu einem Plädoyer für Methodenvielfalt und Multidisziplinarität (Herman, 2003, S. 2). Narrative als Strategien der Sinnstiftung (sensemaking) sind den wissenschaftlichen Erklärungsmethoden in keiner Weise unterlegen. Schon Bruner (1986) unterschied die narrative von der paradigmatischen Denkweise.² Die *Narrative Psychologie* betrachtet das Narrativ als Paradigma für die Sicht, dass Wissen sozial konstruiert ist. Das «narrative Wissen» ist Grundlage für die Theoriekonstruktion und das Theorieverständnis der Narrativen Psychologie (Herman, 2003, S. 2), oder wie es Herman für seine Sicht auf das Narrativ ausdrückt: «Intelligence itself is rooted in narrative ways of knowing, interacting and communicating» (S. 6).

Herman (2003, S. 4–5) integriert strukturalistische, kontextualistische und kognitive Ansätze in eine *integrative Theorie zum Narrativ*:

1. Die *Strukturalisten*, wie Todorov, Genette, Greimas, Barthes plädieren für einen narratologischen Zugang zum der Erzählung zugrunde liegenden *Code* (la langue), der die Generierung und Interpretation der narrativen «Botschaften» ermöglicht (la parole); sie möchten hinter die Stories sehen, auf die Baupläne achten.
2. Die *Kontextualisten* stellen den sozialen Kontext ins Zentrum. Äusserungen sind nicht der Ausdruck präexistierenden Codes, sondern eher ein interaktiv ausgehandeltes Resultat aus Kontext- und Handlungsaspekten (vgl. linguistische Pragmatik). Labov und Waletzky (1967) sind hier einzuordnen; sie haben sowohl linguistische Strukturen als auch interaktionale und kontextuelle Aspekte von Erzählungen untersucht.
3. Die Vertreter des *kognitiven Paradigmas* (Bsp. Schank und Abelson) betonen die Produktionsregeln als Basis einer Story (*story grammar*), unabhängig von variablen Storyinhalten. Erzählungen folgen einem

²narratives vs. paradigmatisches Wissen; bei Bruner: narrativer vs. paradigmatischer Modus.

narrativen Regelsystem und führen semiotische Strukturen mit kognitiven Ressourcen zusammen. Beispiele dafür sind *Scripts*, *Plans*, *Goals* und *Understanding* (Schank & Abelson, 1977).

4. *Integration*: Herman plädiert dafür, dass die semiotischen Strukturen und kognitiven Ressourcen mit einer dritten Komponente trianguliert werden müssen, nämlich mit sozialen Bedingungen und Prozessen.

Herman beschreibt zwei mögliche Strategien für die Rezeption und das Verständnis von narrativen Strukturen:

- Die erste Strategie geht von der Frage aus: «Wie kann man aus Stories Sinn erschliessen?» Welche mentalen Repräsentationen und kognitiven Prozesse ermöglichen uns, Stories zu verstehen? Scripts und Frames können dem Leser/Hörer helfen, Situationen zu verstehen; gleichzeitig aber verändern die neuen Situationen auch diese interpretativen Modelle.
- Die zweite Strategie betrachtet «Stories als Sinnstifter (Stories as sense-making)»: Das Narrativ als Instrument zur Sinnstiftung, als semiotische und kommunikative Ressource, die aktiv angewendet wird und dem Erzähler bei der Realitätsbewältigung hilft (Herman spricht auch von «Stories as tools for thinking») (Herman, 2003, S. 12).

Aus Sicht der kognitiven Linguistik basieren Narrative auf *frames* und *scenarios*. Buchholz (2009) weist auf die Verwandtschaft zum Begriff *Szenisches Verstehen* hin, wie er von Alfred Lorenzer und Hermann Argelander verwendet wird. Lakoff spricht von *deep narratives* (Lakoff, 2009), «narrativen Tiefenstrukturen» auf kulturellem Hintergrund, die grundlegende Muster für unser Denken und unsere Emotionen darstellen.³

Dagegen gibt es eine eher praktisch orientierte Ausrichtung der Erzählanalyse, die in der *Alltagskommunikation* verankert ist. Gülich (2001) spricht von einer *Alltagsmethodologie*, die kognitive und kommunikative Aktivitäten zugleich erklärbar und darstellbar macht (accountable).⁴ Im *Narrativen Interview* (Schütze, 1982, 1983) wird die alltagsweltliche Erzählkompetenz für die Erhebung von Daten und für die Auswertung von autobiographischen Stegreiferzählungen genutzt. Der Erzähler im Narrativen Interview

³Beispiel für ein allgemein gültiges Tiefennarrativ ist die «Vom Tellerwäscher zum Millionär»- Story (the Rags to Riches story), die unsere Sympathien und unser Mitgefühl weckt.

⁴«Account»: Begriff aus der Konversationsanalyse, bezeichnet kommunikative Handlungen der Gesprächsteilnehmer.

ist verschiedenen Regelhaftigkeiten unterworfen, wenn er sich auf das Erzählen einlässt, Begriffe und Konzepte, die bei Schütze entworfen werden, sind Detaillierungszwang, Relevanzfestlegungs- und Kondensierungszwang; dies sind Regeln, auf die sich der Erzähler verpflichtet, wenn er sich auf das Erzählen einlässt, ebenso wie auf das Abschliessen einer einmal begonnenen Struktur (Gestalterschliessungszwang) (Gülich, 2001, S. 1091). Um vergangene Erfahrungen zu rekonstruieren, verwenden Klienten oft *narrative Gesprächsgenres*. Dabei werden die Erfahrungen aus der Vergangenheit von ihrem Kontext losgelöst (dekontextualisiert) und in einem neuen kommunikativen Kontext rekontextualisiert (Günthner, 2005).

Bamberg und Georgakopoulou prägen eine Kategorie Narrativ, die sie als *Small Story* bezeichnen (Bamberg & Georgakopoulou, 2008; Georgakopoulou, 2007) und verorten dieses Narrativ einerseits in der soziolinguistischen Tradition, andererseits vor allem im interaktionalen Paradigma. Die Small Story ist ein besonderer *Kommunikationsmodus* (mode of action); dieser Aspekt ist für die Autorin wichtiger als das Narrativ als Texttyp oder als Produkt. Im Gegensatz zum Narrativ als Ausdruck der grossen autobiographischen Erzählung (Lebensnarrativ) und im Gegensatz zur Text- und Genreorientierung der Erzählforschung stellt Georgakopoulou das *Erzählen als Interaktion* und als *soziale Praxis* ins Zentrum. In dieser interaktionalen Sichtweise werden der Small Story die folgenden Eigenschaften zugeordnet (Georgakopoulou, 2007, S. 4):⁵

1. Das Narrativ ist eingebettet in lokale Gegebenheiten und nicht davon loszulösen; es ist deshalb nur im Kontext zu verstehen.
2. Narrative werden sequentiell abgearbeitet; die Entwicklung und Verarbeitung geschieht laufend in der Interaktion.
3. Narrative sind *emergent*, d. h. sie entstehen in der Interaktion und als Produkt der Verhandlung zwischen den Gesprächspartnern.⁶

⁵vgl. auch Bamberg (2008): «the term <small story> is meant to refer to stories told in interaction; stories that do not necessarily thematize the speaker, definitely not a whole life, but possibly not even events that the speaker has lived through - and now, retrospectively, reflects upon and recounts (often termed <personal stories> or <narratives of personal experience>). Rather, <small stories> are more the kinds of stories we tell in everyday settings (not just research or therapeutic interviews). And these stories are most often about very mundane things and everyday occurrences, often even not particularly interesting or tellable; stories that seem to pop up, not necessarily even recognized as stories, and quickly forgotten; nothing permanent or of particular importance - so it seems.»

⁶Die Autorin verweist auf Hopper (1987), der den Begriff der «emergent grammar» geprägt hat.

4. Narrative sind situativ und lokal verankert: ihre Bedeutung entsteht im Kontext und im Verlauf ihrer Produktion.

Die hier vertretene Sichtweise auf das Narrativ steht im Gegensatz zu einer Auffassung der *Erzählung* als kontextfreie Sprachproduktion, die durch sprachinterne Regeln bestimmt ist. Im Bereich der Grammatik kann dies mit dem Gegensatz zwischen funktionalistischen, auf dem Sprachgebrauch basierenden Theorien und regelbasierten generativen Grammatiktheorien verglichen werden. Der emergente Charakter der Narrativ-Entstehung widerspricht dem Gedanken der *story grammar*.⁷ Erzählen ist ein kontinuierlicher Prozess von erzählender Konstruktion (und Ko-Konstruktion). Auch in Bezug auf den zeitlichen Ablauf und die Evaluation einer Erzählung vertritt Georgakopoulou eigene Positionen. Erzählungen beziehen sich demgemäss nicht a priori auf Vergangenes (Ebene der Zeitlichkeit), sondern sind zeitlich im Erzählkontext verwoben. Die Evaluation einer Erzählung muss nicht als formaler Abschluss gemäss Schema am Ende stehen, sondern Evaluationsschritte werden immer wieder im Gespräch (neu) ausgehandelt.⁸ Georgakopoulou (2007, S. 7) möchte bewusst mit den Regeln und Annahmen der traditionellen textorientierten linguistischen Erzählforschung brechen.

Kern und Quasthoff (2005, S. 20f) beleuchten interaktionale Perspektiven von fantasierten Kindererzählungen und erstellen ein Inventar von Eigenschaften des Narrativs, die sie als *narrative Marker* bezeichnen (narrative devices). Diese Merkmale beschreiben nicht den Ablauf der Erzählung, sondern benennen bestimmte Ereignisse, die das Narrativ charakterisieren und eine besondere Bedeutung haben. So gibt es am Anfang der Erzählung Stellen, an denen ein Thema eingeführt wird oder über ein Thema verhandelt wird (topic proposal, topic ratification). Später kann ein Wechsel des Themas stattfinden (topic shift), die Intonation kann als narratives Kennzeichen verwendet werden, oder syntaktische Besonderheiten wie die Voranstellung einer Konstituente ins Vorfeld (topicalization). Tempuswechsel markieren Aktivitätswechsel; der Wechsel von Vergangenheit zum Präsens markiert einen Aktivitätswechsel in Richtung Interaktivität (shift tense – switch of activity). Ausserdem gibt es verschiedene lexikalische Marker, wie Konjunktionen und Diskursmarker (*try marker* und *misplacement marker*).

⁷Dass Erzählen durch Gesetzmässigkeiten einer *story grammar* gesteuert wird, ist eine Idee aus dem Paradigma der Generativen Grammatik.

⁸Die Erfahrungen mit der Erzählanalyse JAKOB zeigen ebenfalls, dass evaluative Erzählabschnitte nicht immer am Schluss von Erzählungen stehen, sondern häufig auch innerhalb oder sogar am Anfang der Erzählung.

Georgakopoulou (2007, S. 5/6) erwähnt das Problem der *Granularität* unterschiedlicher Erzählanalyseverfahren: Analyseprozeduren auf der Mikroebene untersuchen Erzählsequenzen sehr kleinräumig, subjektiv und einzelfallbezogen (z. B. im Stil der *Konversationsanalyse*), während Verfahren auf der Makroebene Generalisierungen mit Hilfe von «objektiven» Regeln und Gesetzmässigkeiten ermöglichen möchten. Georgakopoulou plädiert dagegen für eine mittlere Ebene, die sie mit *sozialer Praxis* bezeichnet und die sowohl den Einbezug von relativ stabilen, typisierten und verfestigten (prefabricated, vgl. Kap. 4.7) Kommunikationsstrukturen ermöglicht, aber auch dynamische, im Prozess entstehende Kommunikationsaspekte berücksichtigt.

Bamberg und Georgakopoulou (2008) skizzieren eine *Analyseprozedur für small stories* in fünf Schritten, die auf dem Modell der *Positionierung* beruht und eine Analyse ermöglicht, die sich zwischen den Extremen einer mit Details überladenen Mikroanalyse und einer (holzschnittartigen) Makroanalyse bewegt. Das Vorgehen ist zwar auf ein konkretes Beispiel bezogen, lässt sich aber verallgemeinern:⁹

1. Step one: Who are the characters and how are they relationally positioned? (Positioning level 1). Zuerst wird festgestellt, wie die Personen in der Story positioniert sind, sowohl gegenüber einander als auch räumlich und zeitlich (Situativer Kontext der Story).
2. Step two: The interactive accomplishment of narrating? (Positioning level 2). Wie wird die Aktivität des Erzählens interaktiv in diesem Gesprächsausschnitt entwickelt (mit Bezug auf den grösseren Gesprächskontext)?
3. Step three: How is the speaker positioned within the interactive flow of turns that constitute the situation as research? Positioning level 2). Wie positioniert sich der Erzähler im interaktiven Ablauf des ganzen Gesprächs (oder Interviews), was ist die Rolle des Moderators oder Zuhörers und allenfalls der institutionellen Gesprächssituation (wurden z. B. Fragen gestellt, die bestimmte Reaktionen des Erzählers bewirken)?
4. Step four: How is the relation between the four boys managed? (Positioning level 2). Wie entwickelt sich die Beziehung der Gesprächsteilnehmer während der Erzählung (im Beispiel eine Episode zwischen

⁹Die Positionierungsebenen 1 bis 3 werden z. B. bei Lucius-Hoene und Deppermann (2004) beschrieben.

vier Jungen, von denen einer eine Erzählung ankündigt, dann aber aber seine Ankündigung wieder zurückzieht und die Erzählung an einen andern Teilnehmer zu delegieren versucht. Anwesend ist ausserdem der Moderator und mehrere Forschungsassistenten)? Alle Anwesenden in der Gesprächssituation gelten in diesem Fall als Gesprächsteilnehmer.

5. Step five: Who am I in all this? (Positioning level 3). Mit welchen Erzählerrollen skizzieren die Gesprächsteilnehmer die Anderen und sich selbst, wie positionieren sie sich gegenseitig und wie konstruieren sie dabei das eigene Selbst- und Identitätsgefühl?

Das Positionierungskonzept wurde bisher hauptsächlich in der gesprächsanalytischen Erzählforschung beschrieben, so verwenden z. B. Lucius-Hoene und Deppermann (2004) die *Positionierung* und die *Positionierungsanalyse* als Instrument zur empirischen Erforschung *narrativer Identität*. Aktivitäten von Selbst- und Fremdpositionierung sind ein wichtiger Aspekt aller Gespräche und Erzählungen; besonders aber in autobiografischen Erzählungen und in narrativen Interviews spielen sie eine grosse Rolle, weil in diesen Gesprächsformen das Rederecht in der Regel für längere Zeit beim Erzähler bleibt als dies in alltäglichen Gesprächssituationen der Fall ist. Positionierungsaspekte dienen nicht nur der Selbstdarstellung, sondern auch der «Selbsterstellung» und tragen zu den identitätsstiftenden Leistungen des Erzählens bei. Die Positionierungshandlungen spielen ausserdem eine grosse Rolle im sequenziellen Gesprächsablauf; sie entwickeln sich aus dem vorangegangenen Kontext und projizieren gleichzeitig nachfolgende Kontexteigenschaften (vgl. Linell, 2009, und Kap. 6.7).

Günthner und Bückner (2009) beschreiben Konzepte der Selbst- und Fremdpositionierung aus der Perspektive der Konstruktionsgrammatik; die Beiträge des Sammelbandes beschreiben sprachliche Merkmale von Alltagsinteraktionen, die zur Durchführung von Positionierungsaktivitäten eingesetzt werden.

2.2 Das Narrativ in der Psychotherapie

Die Rolle des Narrativs in der Psychotherapie stösst schon seit längerer Zeit auf Interesse (McLeod, 1997). Seit Tonaufnahmegeräte die Aufzeichnung therapeutischer Gespräche ermöglichen, lassen sich Narrative von Patienten linguistisch und gesprächsanalytisch untersuchen. Narrative Redemuster

und narrative Darstellungsstrategien zeigen, dass Erzählen eine Konstruktionsleistung im Dienst der Selbstvergewisserung und der sozialen Resonanz, nicht aber eine beobachtungstreue Dokumentationsleistung ist (Boothe, Grimm, Hermann & Luder, 2010). Die *Narrative Psychologie* (Sarbin, 1986) betrachtet das Narrativ als grundlegend für die menschliche Erfahrungsorganisation. Dass Menschen ihre Erfahrungen in Geschichten verwandeln und in Form von Erzählungen kommunizieren, bildet ihren Ausgangspunkt. Die mit dem Erzählen verbundenen Prozesse des *meaning making*¹⁰ (Bruner, 1990, 2004) führen zur Ausbildung der *narrativen Identität*.

Luborsky und Crits-Christoph (1990) veröffentlichten 1990 das *CCRT* (Core Conflictual Relationship Theme Method)¹¹ als Instrument für die Analyse von *Beziehungsepisoden*, die ebenfalls unter den Begriff der *Erzählung* subsumiert werden können. Das CCRT wurde in sehr vielen Forschungsprojekten angewendet und ist dementsprechend gut dokumentiert (für eine neuere Zusammenfassung siehe Albani, Pokorny, Blaser & Kächele, 2008).

Narrative aus Therapiegesprächen als Darstellung persönlichen Erlebens werden häufig im Rahmen *konversationsanalytischer Studien* untersucht (siehe auch Kap. 3.2.1). Erzählungen werden in diesem Kontext als *personal narrative* bezeichnet: «We consider a personal narrative to be any segment of talk in which clients build an account of their life events with a temporal and causal order» (Ochs & Capps, 2001). Auf konversationsanalytischem Hintergrund untersuchen Bercelli, Rossano und Viaro (2008) Patientenerzählungen in Bezug auf die Funktionen und Aufgaben, die diesen im Laufe des Therapiegesprächs zukommen und die je nach Gesprächsphase anders aussehen. Sie bestimmen zwei unterschiedliche Phasen im Verlauf einer therapeutischen Sitzung. Erzählungen, die in einer *Explorationsphase* der Sitzung präsentiert werden, bringen vor allem weitere Informationen für den Explorationsablauf zutage. In *Elaborationsphasen* der Sitzung werden sie von den Klienten meistens als Antworten und Präzisierungen auf Deutungen des Therapeuten eingebracht und zeigen, wie der Klient die Intervention des Therapeuten verstanden hat.

In einer Übersicht zu neueren erzählanalytischen Forschungsverfahren verweisen Georgaca und Avdi (2009) auf die prominente Rolle, die das Narrativ/die Erzählung bei der *Darstellung der Lebensgeschichte* spielt (Angus & McLeod, 2004). Das Narrativ ist der Ausdruck und die Reflexion individueller psychologischer Prozesse.¹² Die Erzählanalyse untersucht nur

¹⁰vgl. «meaning construction» und «Bedeutungskonstitution», siehe Kap. (3.3).

¹¹deutsch: ZBKT, Zentrales Beziehungskonflikt-Thema.

¹²Die Autorinnen stellen auch die *Erzählanalyse JAKOB* vor als eine «Public mise en

die Narrative des Patienten, nicht den gesamten Therapiekontext und alle Beteiligten, so auch das «Narrative Processes Coding System» genannte Verfahren (z. B. Hardtke, Levitt & Angus, 2003), das entwickelt wurde, um Strategien und Prozesse im Therapiegespräch zu identifizieren und in eine Darstellung zu bringen, die die Erfahrungen des Klienten und seiner Mitmenschen sowohl organisiert als auch repräsentiert:

Diese Darstellung entsteht sowohl innerhalb einer Therapiesitzung in Form von Mikronarrativen als auch im Verlauf einer gesamten Therapie in Form des [...] makronarrativen Rahmens, der die singulären Ereignisse, die in der therapeutischen Beziehung beschrieben werden, aufeinander bezieht und vernetzt. Die Arbeit entwickelt zunächst eine Narrative Prozesstheorie der Psychotherapie (Narrative Processes Theory of Therapy) und beschreibt dann das narrative Kodiersystem (Narrative Processes Coding System), das entwickelt wurde, um empirische Schlüsselkomponenten des Modells identifizieren und evaluieren zu können. (Hardtke et al., 2003)

Die *Konversationsanalyse* dagegen untersucht interaktionale Prozesse. Der Begriff der *formulation* (eine Art Zusammenfassung der Gesprächsstandpunkte) ist in diesem Kontext wichtig, wie Georgaca betont. *Formulations* scheinen zwar auf den ersten Blick blossе Zusammenfassungen, die Gesprächsteilnehmer setzen aber durch ihre selektive Wahrnehmung Präferenzen für die folgenden Turns (vgl. das Zitat auf S. 136. «formulations have been the royal road into the practices of psychotherapy» (Antaki, Barnes & Leudar, 2005, zit. nach Georgaca & Avdi, 2009, S. 239).

2.3 Die Erzählanalyse JAKOB

Ein wesentliches Merkmal psychodynamischer Behandlung ist die «Aneignung persönlicher Geschichte» des Patienten als «Modellierung und Um-Modellierung von Erfahrung» (Boothe, 2004). Wiederkehrende autobiografische Darstellungsmuster können als Bestandteile von Erzählungen, aber auch in der kommunikativen Inszenierung erkannt und aufeinander bezogen bearbeitet werden. Erzählen, Inszenieren, Erinnern, Alltagspraxis: die wechselseitigen Bezüge kommen im Therapieprozess zur Anschauung und

scene of psychodynamic conflicts. A means of assessing the client's underlying pathology» (Georgaca & Avdi, 2009, S. 235).

zur Reflexion. Therapeut und Patient können die Erzählbedingungen den Bedingungen der Lebenspraxis gegenüberstellen (Boothe et al., 2010).

2.3.1 Modellierungsfunktionen

Erzählen bezieht sich zwar auf Begebenheiten der Vergangenheit, ist aber auch ein Mittel sprachlicher Inszenierung in der Gegenwart, der Erzähler führt Regie und erzeugt Spannung, indem die ZuhörerIn auf eine spezifische Art von einer Ausgangssituation zu einer Ergebnissituation geführt wird. In der Schilderung von Taten und Begebenheiten geben Erzähler auf dramatische Weise zu erkennen, worum es ihnen geht. Dabei modelliert das Erzählen das Erlebte auf vierfache Weise (Arboleda et al., 2010):

Aktualisierung: Erzählen aktualisiert das Vergangene. Es stellt eine Verbindung zwischen Vergangenem und gegenwärtiger Situation her und macht damit klar, dass etwas aus der Vergangenheit Erzähltes auch im Hier-und-Jetzt unter anderen Bedingungen aktuell und wichtig ist.

Soziale Integration: Erzählen hat eine soziale Integrationsfunktion. Es gibt der erzählenden Person eine Identität vor dem Zuhörenden, eine individuelle Bedeutung, stellt sie in einen sozialen Kontext und gibt ihr eine persönliche Geschichte. Allerdings gelingt dies nur dann, wenn die ErzählerIn den Zuhörer emotional in die Erzählung verwickeln kann, wenn dieser ihre Perspektive übernimmt und ihr ein bestätigendes Echo gibt.

Reorganisation: Erzählen reorganisiert das Erlebte. Es verarbeitet und bewältigt Angst und andere Formen von Destabilisierung und Desintegration und ermöglicht einen Gewinn an Kontrolle und Sicherheit in konflikthaftem Erleben.

Restitution oder Wunscherfüllung: Erzählen restituiert das Erlebte. Es reinszeniert das Vergangene sprachlich neu in einer Version, die man für wirklich erlebt und für wahr hält. Implizit wird damit auf den Wunsch verwiesen, wie das Erlebte hätte ausgehen sollen. Der Wunsch kann durch das Erzählen nachträglich erfüllt oder das erzählte Ereignis zumindest in Richtung Wunscherfüllung neu inszeniert werden. Es ist das Wunschemoment, welches die Spannung einer Erzählung erzeugt und antreibt.

Im therapeutischen Dialog werden die auf der Probebühne der Erzählung eingenommenen Rollen reflektiert und exploriert und zum szenischen

Geschehen der therapeutischen Situation in Beziehung gesetzt. Neue Beziehungserfahrungen des Patienten sollten auch in veränderten Beziehungsnarrativen zum Ausdruck kommen, ebenso sollten narrative Modelle der therapeutischen Kooperation bis zum Ende der Behandlung eine veränderte Organisation aufweisen.

2.3.2 Anwendung

Die *Erzählanalyse JAKOB* (Arboleda et al., 2010; Boothe, 2004; Boothe et al., 2010, 2002) ist ein qualitatives Untersuchungsinstrument für *Alltags-erzählungen*, die in kompakter Form emotional bedeutsames Selbst- und Beziehungserleben offenbaren. Die Erzählanalyse JAKOB ermöglicht narrative Einzelfallanalysen auf dem Hintergrund von psychodynamischen und psychoanalytischen Theorien. Die Therapiestunden werden mit dem Einverständnis der Patienten aufgezeichnet und transkribiert, und die Erzählungen werden nach bestimmten Regeln extrahiert. Erzählungen sind episodische Handlungsabläufe, die räumlich und zeitlich fixiert sind und einen Spannungsbogen mit Start, Entwicklung und Ergebnis gleich einem Bühnengeschehen aufbauen. Der Erzähler als Regisseur statet die Szene mit Requisiten und Kulissen aus und inszeniert seine eigenen und die Aktionen der Mitakteure; er führt uns einerseits in der Rolle des Erzählers eine Szene vor, daneben stellt er sich selbst als Akteur in dieser Szene dar. Für die weitere Verarbeitung wird der extrahierte Erzähltext manuell segmentiert (Subjekt-Prädikat-Verbindungen). Ein Segment wird in der Erzählanalyse JAKOB als *Frame* bezeichnet, ein Frame ist eine Akteur-Aktion-Objekt-Struktur, die vier darin enthaltenen *Slots* beantworten die Frage: «wer tut was wie?» (Akteur – Aktion – Patiens – Instrument) oder «Was geschieht in Bezug auf wen/was wie?» Die *linguistische Analyse*¹³ dieser Segmentstruktur erfolgt in mehreren Schritten und verarbeitet morphologische, syntaktische und semantische Merkmale der Textsegmente mit Computerunterstützung. Das in der Erzählung verwendete Vokabular wird darauf mit Hilfe eines vorgegebenen, psychodynamisch/psychoanalytisch orientierten Kategoriensystems kodiert, um die Dramaturgie der Erzählung und ihre Vermittlung durch den Erzähler durch präzise *lexikalische Analyse* zu ergründen. Das Kategoriensystem umfasst Codes für das Personal, die Requisiten, Kulissen und den Bühnenraum, aber vor allem für die Aktionen (Verben), d. h. für die sprachlichen Handlungen von Erzähler und Erzählfiguren.

¹³Dies gilt für die computerunterstützte Version der Erzählanalyse JAKOB, die mit Hilfe der Internetapplikation *AutoJAKOB* durchgeführt wird.

Die sequentielle Abfolge der Segmente wird zusätzlich strukturiert durch die *Regie* oder Erzählregie, die nähere Bestimmung der Segmente als *episodisch* (Kernsegment) oder *nicht episodisch* (Rahmensegmente). *Kernsegmente* stellen die eigentliche Handlung in der erzählten Welt mit der Darstellung der Ich-Figur dar, während die *Rahmensegmente* Kommentare, Bewertungen oder andere Informationen (der Erzählerin) ausserhalb der erzählten Welt darstellen. Zur Regie gehört ausserdem die Einteilung der Segmente in Start-, Entwicklungs- und Ergebnisphase, die idealerweise in dieser Abfolge in der Erzählung identifizierbar sind. Die Operationalisierung dieser Strukturen ist in der Kurzanleitung zur Erzählanalyse JAKOB dokumentiert (Arboleda et al., 2010). Auf Grund des Figurenrepertoires (der Akteure) und der Codes in der Erzählung werden unter dem Oberbegriff der *Sozialen Integration* Auswertungsschritte bezüglich der Positionierung, der Handlungsinitiative und der Darstellung der Figuren durchgeführt. Das *Akteurschicksal* beschreibt die Handlungsinitiative der Ich-Figur, *Zentrierung/Marginalisierung* beschreibt die Positionierung der Ich-Figur gegenüber den andern Akteuren, und die *Präsentation* der Ich-Figur und der andern Akteure in der Erzählung wird auf den Achsen *Macht*, *Nähe* und *Autonomie* erschlossen.

Der erste Handlungsimpuls und die Setzung, weitere Gestaltung und Entwicklung des dramaturgischen Potenzials im Spannungsbogen der Erzählung erhalten besondere Beachtung und werden als *Spielregel* bezeichnet. Die Spielregel besteht in der Formulierung der *Startdynamik* der Erzählung und davon abgeleitet der Formulierung von *SOLL* und *ANTISOLL* als hypothetische Szenarien der optimalen Entwicklung oder der Katastrophe in der weiteren Entwicklung der Erzählung. Diese Szenarien dienen der Erschliessung der nachfolgenden Analyseschritte der Konfliktodynamik und werden der tatsächlichen Entwicklung (SEIN) in der Erzählung gegenüber gestellt.

Nach der textnahen Analyse der *Erzähldynamik* werden im interpretativen Auswertungsteil Hypothesen zur *Konfliktodynamik* gebildet, die als Kompromissbildung aus dem Zusammenspiel von prototypischen *Wunschthemen*, *Angstmotiven* und *Abwehrbewegungen* verstanden wird. Ziel der Analyse ist neben den erzähl- und gesprächsanalytischen Erkenntnissen eine psychodynamische Konflikt-, Beziehungs- und Prozessdiagnostik.

2.3.3 Erzählung und Dramaturgie

Patienten erzählen in der Therapie. Sie fokussieren damit auf bestimmte Themen. Hier das Beispiel einer prototypischen Erzählung der Patientin

Amalie (2):

(2) Erzählung Amalie: «Ein blöder Studienrat»¹⁴

1 im Konzert war ich gestern Abend
 2 und da sass ich auch neben so einem blöden Studienrat
 3 der mich genau kennt
 4 erstens macht er Sport in unserer Schule, kommt da zu uns ins
 Lehrerzimmer, grüsst und
 5 und in einer Arbeitsgemeinschaft war er neulich mit mir auch
 6 und gestern Abend da hat er ringsum seine Flossen hingestreckt
 7 und ich sass neben ihm, zufällig
 8 aber der hat getan
 9 als kenne er mich nicht
 10 es hat mich dann so geärgert
 11 dass ich mich nach der Pause einfach weggesetzt habe
 12 es war sowieso so eng da vorne
 13 ich habe mich dann so furchtbar geärgert
 14 dass das ganze Konzert futsch war
 15 ich konnte nicht mehr
 16 bin nach Hause Musik auflegen
 17 hab ich mehr davon

Die Erzählanalyse JAKOB stellt die *Erzählung* als Untersuchungseinheit ins Zentrum und analysiert diese in einem ersten Schritt relativ kontextfrei, um die Unvoreingenommenheit des Untersuchers bei der Analyse der Dramaturgie und der Erzähldynamik zu bewahren. In einer späteren Analysephase oder nach der Analyse mehrerer Erzählungen werden auch *Kontextfaktoren* einbezogen¹⁵.

Die *Bühnenmetapher* ist wichtig für das Selbstverständnis der Erzählanalyse JAKOB. Das Verfahren bedient sich einer Bühnenterminologie, welche auf dem *psychoanalytischen Bühnenmodell* von Thomä und Kächele (2006) beruht. Die Erzählanalyse JAKOB fragt, wie der Erzähler als Regisseur sein Anliegen narrativ inszeniert und versucht, das Szenario herauszuarbeiten, das der Erzähler im Medium der Sprache herstellt, wenn er den Raum des Geschehens bestimmt, seine Figuren ausstaffiert und ihren Aktionsradius festlegt (Arboleda et al., 2010). Die *Theatermetapher* in Bezug auf das Erzählen wird auch von Günthner (2008) verwendet. Sie verbindet die Aspekte der Theatralität von Alltagserzählungen mit dem Anliegen einer Grammatik der Gesprochenen Sprache (siehe Kap.3.2.2). Grammatische Strukturen und Muster werden in der Interaktion konstruiert, verfestigen sich im Laufe

¹⁴Die Erzählung ist segmentiert, die Segmente sind nummeriert.

¹⁵Vgl. dazu die oben erwähnte Forderung von Georgakopoulou (2007), dass Narrative nur im Kontext zu verstehen sind.

der Zeit, können aber auch wieder modifiziert werden. Diese grammatischen Phänomene werden als Mittel der Inszenierung, als *Theatralisierungsverfahren* in Alltagserzählungen eingesetzt (S. 3). Darunter fallen beispielsweise die Auswahl von direkter oder indirekter Rede bei der Präsentation der Erzählung, die Verwendung des Präsens, die Verbspitzenstellung oder die Verwendung von Infinitkonstruktionen (Sätze ohne finites Verb) als Mittel der szenischen Performanz. Goffman (2008) hat Erzählen als Inszenierung, Alltagserzählungen als szenische Darbietungen beschrieben. Der Erzähler führt dem Publikum ein kleines Drama vor; der Fokus liegt dabei auf der Art und Weise der Darbietung, nicht auf der zu vermittelnden Information. Aufführung und Alltagserzählung sind beide weder fixier- noch tradierbar, flüchtig und transitorisch (Günthner, 2008, S. 5). Die szenische Rekonstruktion erfolgt mit einer *Rekontextualisierung*, die Kommunikationsabsichten des Erzählers verändern die ursprüngliche Szene, «durch die Stimmen der Figuren dringt die wertende Stimme des Erzählers» (S. 8). *Figurentext* und *Erzähltext* überlagern sich.

Erzähltypen

Die prototypische Form der Erzählung, die in der Erzählanalyse JAKOB untersucht wird, ist die *episodische Erzählung* (Bsp. 2). Diese kommt jedoch in ihrer «reinsten» Ausprägung nicht so oft vor; viel häufiger sind Erzählformen, in denen episodische Darstellung vermischt ist mit Reflexion, Argumentation oder mit der Schilderung von immer wiederkehrenden Situationen. Was gibt es neben der episodischen Erzählung für weitere Erzählformen? In der Psychotherapie kommen vor allem drei Erzählformen immer wieder vor (Boothe, 2009):

- Die *episodische Erzählung* hat einen räumlich und zeitlich fixierten episodischen Handlungsablauf vom Start über die Entwicklung zum Ergebnis und wird häufig durch Zeitmarkierung («Am letzten Sonntag bin ich...») eingeleitet.
- Die *iterative Erzählung* beschreibt ein Ablaufmuster, das immer wieder vorkommt und das sich nicht auf eine bestimmte Episode bezieht (Erzählanfang z. B. durch «Immer wenn ich...»).
- Die *Karriereerzählung* schildert einen lang dauernden Prozess oder einen Ablauf; häufig ist dies die Beschreibung eines Krankheitsverlaufes.

2.3.4 Kodierung und Lexikon

Das in der Erzählung verwendete Vokabular wird mit Computerhilfe auf Grundlage des JAKOB-Lexikons kodiert. Ermöglicht und erleichtert wird die computergestützte *Kodierung* durch die vorangegangene Segmentierung (Bestimmung der Satzgrenzen),¹⁶ durch eine genaue Transkription der Gesprächstexte (keine orthographischen Fehler) und durch die vorgängige linguistische Bearbeitung der elektronischen Texte mit Computerhilfe (Lemmatisierung, partielles Parsing mit Bestimmung der Konstituentenstruktur). Diese Voraussetzungen ermöglichen es, Codes aus dem Lexikon einzelnen Ausdrücken oder Wörtern zuzuordnen. Eine manuelle Nachbearbeitung und Plausibilitätskontrolle ist aber immer nötig, vor allem auch deshalb, weil viele der gesuchten Ausdrücke nicht im Lexikon vorkommen.

Die *Kodierung* soll die weiteren Schritte der Erzählanalyse unterstützen und Interpretationshilfen geben. So dienen die Codes beispielsweise dazu, die *Selbstpräsentation* der Erzählerin gegenüber ihrer *Präsentation der Ich-Figur* zu charakterisieren und die soziale Involviertheit der Akteure zu erschliessen (*Macht*, *Nähe* und *Autonomie*). Die Verbcodes der Segmente der *Startdynamik* dienen als Grundlage für die hypothetische Formulierung der *Spielregel* für den weiteren Verlauf der Erzählung. Die Spielregel wiederum dient als Ausgangspunkt für Bestimmung der *Wünsche*, *Ängste* und *Abwehrmechanismen*, die in der Erzählung dargestellt werden.

Der Kodierung der *Verben* kommt in der Erzählanalyse JAKOB eine besondere Bedeutung zu; die Verben repräsentieren die Aktionen, mit denen die Erzählerin die Szene dramaturgisch gestaltet. Aus den *lexikalischen Wahlen* sollen psychodynamische und für die Interpretation der Erzählung relevante Kategorien abgeleitet werden können. Die Kategorien sind im JAKOB-Kodiersystem festgehalten und dokumentiert (siehe Kapitel 2.3.4); die Kodierung von Wörtern und Ausdrücken mit diesen Kategorien geschieht mit Hilfe des JAKOB-Lexikons (ca. 7000 Einträge). Falls die zu kodierenden Ausdrücke nicht im Lexikon vorkommen, muss die Kodierung manuell durchgeführt werden. Der praktische Aufbau des JAKOB-Lexikons wird im Kap. (7.1) besprochen.

¹⁶Ob die Einteilung in Segmente als Propositionen mit Prädikat-Argumentstruktur den Eigenarten des Gesprächs gerecht wird, ist umstritten (Auer, 2010); für die Bearbeitung mit NLP-Werkzeugen bietet sich jedoch keine Alternative an.

Das Kodiersystem

Mit der dramaturgischen Kodierung wird die *lexikalische Wahl* des Erzählers erfasst. Dazu wird ein auf psychoanalytischen Konzepten beruhendes Kategoriensystem verwendet. Die «Lexikalisierung schafft in der narrativen <Welt-für-mich> den erzähldynamischen Erwartungs- und Erfüllungsraum» (Boothe, 2011, S. 60). Die lexikalischen Setzungen sind für die Darstellung der Situation, für die Plot-Organisation und für die Erzähldynamik von zentraler Bedeutung.

Der theoretische Hintergrund der Kodierung wird im Manual der Erzählanalyse JAKOB (Boothe et al., 2002) und im Kodiermanual (Boothe, 2002) beschrieben (siehe auch Boothe, 2011). Für die praktische Anwendung der dramaturgischen Kodierung steht das JAKOB-Lexikon zur Verfügung, in der computerunterstützten Anwendung der Erzählanalyse wird das Vokabular der Erzählung automatisch mit Hilfe dieses Lexikons kodiert. Für die manuelle Kodierung oder für die Kodierung der noch nicht im Lexikon enthaltenen Wörter und Ausdrücke kann das JAKOB-Lexikon durchsucht werden¹⁷, zudem steht ein spezielles Kodiermanual zur Verfügung, das die einzelnen Kategorien beschreibt und mit Ankerbeispielen verdeutlicht (Boothe, 2002).¹⁸

Das JAKOB Kodiersystem ist in der Lexikon-Datenbank enthalten und unterscheidet die folgenden Kategorien:

- *Dimensionen*: Hier sind die *fünf Dimensionen* der Verbkodierung gemäss JAKOB-Manual erfasst (Tabelle 2.1).
- *Verbcodes*: umfasst die 93 Kodiermöglichkeiten für Verben (die in den fünf oben genannten Dimensionen untergebracht sind).
- *Klassen*: Unter diesem Oberbegriff werden die folgenden JAKOB-Kategorien aus dem Manual zusammengefasst: Positionen, Richtungen, Konstellationen, Figuren, Dinge, Modifikationen, Vergleiche, Qualifizierungen und Trennzeichen.

Die *Verbcodes* sind die zentralen Elemente im Codesystem; die differenzierte Kodierung der Verben entspricht dem Stellenwert der *verbalen Handlungen* in der Erzählanalyse JAKOB: Die Segmente der Erzählung werden daraufhin befragt, wie die *Akteure handeln* («wer tut was wie?» oder «was geschieht wem wie?»). Innerhalb der Dimensionen sind die 93 Verbcodes

¹⁷URL: <http://www.jakob.uzh.ch/lexikon/>

¹⁸URL: <http://www.jakob.uzh.ch/lexikon/> > Erläuterung der Verbcodes und Dimensionen.

zusätzlich in *Subdimensionen* eingeteilt. Tabelle (2.1) zeigt die Dimensionen und Subdimensionen der Verbcodes, die vollständige Liste der Codes ist in der JAKOB-Kurzanleitung enthalten (Arboleda et al., 2010).¹⁹

Tabelle 2.1: Die Verbkodierung in der Erzählanalyse JAKOB

Dimension	Subdimensionen	Codes
Geschehen	Markieren, Existieren, Vorgehen	30 Codes, zB. ANF (anfangen), ATT (angreifen)
Fühlen	Annähern, Distanzieren	6 Codes, z. B. WUN (bewundern), ACH (verachten)
Wollen	Intendieren, Festlegen	21 Codes, z. B. RES (resignieren), WIL (wollen)
Handeln	Präsentieren, Interagieren	18 Codes, z. B. DAR (darstellen), KAM (kämpfen)
Schaffen	Ordnen, Kontrollieren, Binden	18 Codes, z. B. ORD (ordnen), KON (kontrollieren)

Stellvertretend für das ganze Kodiersystem sollen hier die fünf Dimensionen aus der Tabelle (2.1) charakterisiert werden (nach Boothe, 2011):

- Die Dimension *Geschehen* legt den Schwerpunkt auf körperliche Vorgänge und beinhaltet Verben der Zuständlichkeit, der Bewegung, der Prozessualität, des Verrichtens: «Verwendet ein Erzähler Verben dieser Dimension, wie z. B. laufen, klopfen, sprechen, aber auch stehen, verharren, so zeichnet er eine Figur als sich bewegend, tätig, mit (körperlichen) Verrichtungen beschäftigt aus. Es handelt sich grundsätzlich um episodisches Geschehen, häufig erfassbar als physisch zu beobachtender Ablauf, sowohl in der belebten als auch der unbelebten Welt.» (Boothe, 2011, S. 62)
- Die Dimension *Fühlen* setzt den Schwerpunkt auf Emotionen und beinhaltet Verben der Appetenz: «Verwendet ein Erzähler Verben dieser Dimension, wie z. B. lieben, hassen, bewundern, verachten, so zeichnet er eine Figur durch ein Angezogen- oder Abgestossensein bezüglich eines Gegenstands seiner Aufmerksamkeit aus. Die Verwendung von Verben der Appetenz im Narrativ lässt die entsprechend gekennzeichnete Figur in einer objektgerichteten Neigung auftreten; [...]» (S. 63)

¹⁹URL: <http://www.jakob.uzh.ch/lexikon/> > JAKOB Kodiertabelle/Liste der Aktionscodes.

- Die Dimension *Wollen*, mit Schwerpunkt auf Motivation, beinhaltet Verben der Intentionalität: «Verwendet ein Erzähler Verben dieser Dimension, z. B. wollen, wünschen, resignieren, aber auch ablehnen, verabscheuen, so zeichnet er eine Figur als intentional orientiert aus. Die Figur bezieht hinsichtlich eines Gegenstandes des Interesses einen Standpunkt oder ringt um einen Standpunkt (wie etwa bei zweifeln).» (S. 64)
- Die Dimension *Handeln* legt den Schwerpunkt auf Handlung und beinhaltet Verben der intelligenten, regelgeleiteten Interaktion: «Verwendet ein Erzähler Verben dieser Dimension, z. B. spielen, arbeiten, kämpfen, so zeichnet er eine Figur als intelligent handelnd aus, als Mitglied einer sozialen Gemeinschaft mit spezifischen Verkehrsformen und Regelungen des Zusammenlebens.» (S. 65)
- Die Dimension Schaffen bezieht sich auf zielgerichtete Beziehungsgestaltung und beinhaltet Verben der Beziehungsfiguren: «Verwendet ein Erzähler Verben dieser Dimension, z. B. kontrollieren, scheitern, verführen, abgrenzen, so zeichnet er eine Figur im Rahmen einer Beziehungskonstellation. Beziehungskonstellation meint dabei nicht nur Interpersonalität, sondern auch eine institutionelle Beziehungskonstellation oder Sachbezüge.» (S. 67)

2.4 Lexical Choice

Besonderer Wert wird innerhalb der Erzählanalyse JAKOB auf die *lexikalischen Wahlen* gelegt. Dieser Ansatz geht auf Donald Spence zurück, der innerhalb der Psychotherapieforschung auf das Phänomen *lexical choice* schon zu Beginn der achtziger Jahre aufmerksam machte (Spence, 1980). Ausgehend von Freuds Gedanken zu Versprechern und andern sprachlichen Fehlleistungen (Freud, 1904) macht sich Spence Gedanken über die Funktion der lexikalischen Setzungen, wie sie im Therapiegespräch von Patient und Therapeut verwendet werden. Analog zu den Versprechern findet Spence auch im Therapiegespräch Wortwahlen, die verborgene Intentionen und Motive des Patienten wiedergeben. Spence spricht von einem «lexikalischen Leck» (lexical leakage); das Unbewusste sickert durch den Sprachgebrauch hindurch. Wichtige latente Themen des Gesprächs können durch die spezifische Wortwahl des Patienten offenbart werden. Als Beispiel nennt er die *Metapher*, die seiner Ansicht nach als Vehikel für den Ausdruck von latenten Themen dient. Konventionelle Metaphern stellen Bedeutung auf

der Ebene des ganzen Metapherausdrucks her; die Bedeutung des metaphorischen Ausdrucks wird vom Hörer sofort verstanden, die lexikalischen Details werden dadurch nicht mehr beachtet und können so eine zusätzliche verborgene Bedeutung transportieren. Metaphern erlauben es, spezifische Wörter in einem «harmlosen» Kontext auszudrücken (das Gleiche gilt auch für Zitate und Sprichwörter). Dieser «Übersetzungsprozess» kann sowohl im Sprecher als auch im Hörer ablaufen.

Aus drei Gründen werden aber gemäss Spence (1980, S. 129) Sprecher und Hörer daran gehindert, den latenten Inhalt zu verstehen:

1. Der mündliche Diskurs läuft schnell ab, so dass weder Sprecher noch Hörer Buch führen können über die einzelnen verwendeten Wörter; es wird deshalb keine bewusste lexikalische Analyse vorgenommen (Spence nennt als Beispiel die Einführung des Themas ‹Tod› durch die Metapher «sich zu Tode langweilen», deren figurative Bedeutung nicht den Tod betrifft. Sprecher und Hörer fokussieren auf die figurative Bedeutung der Äusserung und haben deshalb gar keine Zeit, sich detailliert über die Wortwahl Gedanken zu machen.
2. Noch schwieriger ist es für Sprecher und Zuhörer, Buch zu führen über die wiederholte Verwendung und die Abwandlungen von bestimmten Ausdrücken und Wörtern.
3. Der latent übermittelte Inhalt ist im *isolierten Einzelwort* enthalten, während wir, um Sprache zu verstehen, das Wort als Bedeutungsträger immer in einem bestimmten Ausdrucks-, Satz und Diskurskontext verorten.

Trotzdem ist die Wortwahl nicht zufällig. Spence spricht von *underground code*: Themen können versteckt zur Sprache gebracht werden, Äusserungskomponenten mit sonst konventionalisierter Bedeutung erhalten plötzlich eine andere, individuelle Bedeutung.²⁰ Dank ihrer konnotativen Eigenschaften geben Metaphern und figurative Sprache dem Patienten die Möglichkeit etwas mitzuteilen, ohne das Thema explizit zu formulieren.

Peräkylä (2004) untersucht mit konversationsanalytischen Methoden interaktionale Aspekte der psychoanalytischen Deutung und thematisiert das Phänomen des *lexical choice* speziell auf den Therapeuten bezogen. Dieser interpretiert und verknüpft die Erfahrungen/Erzählungen des Patienten aus

²⁰Das umgekehrte Phänomen findet man in der formalen Denkstörung des *Konkretismus* (Symptom der Schizophrenie), bei der an der konkreten Wortbedeutung festgehalten wird (Luif, 2006).

der Vergangenheit mit den neuen Erfahrungen im Hier und Jetzt, die den Patienten die Therapie aufsuchen lassen, durch Parallelen in der Wortwahl (Konstruktionswahl), so dass es dem Patienten möglich wird, vergangene und aktuelle Erfahrungen zu verlinken. Mit Hilfe der Gesprächsanalyse zeigt er die interaktionalen Praktiken auf, mit denen die Verlinkung bewerkstelligt wird: «interpretations propose *pattern matches* between current and past events...» (Spence, 1984, S. 145).

Mit der individuellen Wortwahl können also Patienten (oder Sprecher) verschiedene Bedeutungen und Mitbedeutungen transportieren, die beim Gegenüber verstanden werden oder auch nicht. Teilweise hängt damit die umstrittene Frage zusammen, ob es eine *wörtliche Bedeutung* (literal meaning) gibt und was diese während der Kommunikation bewirkt. Speziell für eine Methode wie die Erzählanalyse JAKOB, mit der Kodierung des Vokabulars und besonders der Verben, stellt sich diese Frage. Gibt es eine situations- und kontextunabhängige Wortbedeutung, die im Lexikon verzeichnet (und dementsprechend auch kodiert) werden kann und die quasi als Grundbedeutung beim Verstehen eingesetzt wird, oder ist für das Erschliessen der Wortbedeutung immer pragmatisches und situatives Kontext- und Weltwissen erforderlich (vgl. Deppermann & Spranz-Fogasy, 2001, S. 1157)?

Es gibt viele Experimente, die aus psycholinguistischer Sicht die Frage zu klären versuchen, ob beim Sprachverstehen zuerst die wörtliche Bedeutung (literal meaning) von Wörtern und Ausdrücken (im mentalen Lexikon) analysiert werden muss oder ob die Verarbeitung direkt die figurative Bedeutung abrufen. Experimentelle Ergebnisse sprechen für die letztere Variante (Gibbs, 1989, 1994). Gibbs Studien zeigen, dass die Verarbeitung figurativer Ausdrücke erfolgen kann ohne eine vorgängige Analyse der wörtlichen Bedeutung. Die Verwendung figurativer Sprache entspringt gemäss Gibbs dem Wunsch, verschiedene weitere Aspekte neben der propositionalen Bedeutung zu kommunizieren (vgl. die Aussagen von Spence oben). Metaphorische und metonymische Aussagen bewirken eine affektive Unmittelbarkeit, eine zusätzliche Dimension der Interaktion (Gibbs, 1989, S. 247–250).

Die Frage des *literal meaning* wird im Kapitel zur Phraseologie (4) wieder aufgegriffen werden.

2.5 Text und Verknüpfung

Hermann Argelander entwirft in seinem Buch «Der Text und seine Verknüpfungen» interessante Aspekte des psychoanalytischen *Textverstehens* (Argelander, 1991). Die psychoanalytische Methode beschäftigt sich immer

wieder mit dem Sprechen und mit der Sprache, ihren Strukturen, ihren Subtexten, ihren Lücken und ihren Verknüpfungen. Eine Deutung schliesst eine Zusammenhangslücke und thematisiert dadurch etwas Neues, was inhaltlich nicht im Text ist. Brüche, Lücken und Zusammenhangslücken zeigen, dass der Text deutungs- oder interpretationsbedürftig ist.

Argelander begann, die *Textverknüpfungen* zu formalisieren, um damit die Lücken und Brüche besser sichtbar zu machen. Grundlage der Verknüpfungen sind die einfachen *Aussagen* oder Sätze eines Textes. Er unterscheidet zwischen eindeutigen und widersprechenden Verknüpfungen. *Sinnvariationen* ergeben sich durch Verknüpfungsänderungen. Daraus entsteht eine komplizierte Textverknüpfungsstruktur mit Knoten, Abzweigungen und Rückbezügen. Die Verknüpfungsstruktur zeigt die Suche nach *Sinn und Bedeutung* als dem zentralen Element der psychoanalytischen Forschung. Aussagen werden in einem sequenziellen Vorgehen in einen Zusammenhang gestellt. Ein einfacher Satz ist auch schon eine Zusammenhangsbildung. Manchmal kann ein Sachverhalt nicht mit dem vorherigen verknüpft werden, der Text wird unterbrochen, fragmentiert, und es beginnt ein neuer Textabschnitt, Ausdruck einer Textkompression.

Die nachfolgende Erklärung und komplizierte Anwendung des Kodiersystems, die bei Argelander breiten Raum einnimmt, bringt allerdings einstweilen für das vorliegende Projekt keine neuen Erkenntnisse.

2.6 Fazit zu Narrativ und Erzählanalyse

Das *Narrativ* oder die *Erzählung* stellt ein geeignetes und interessantes Format für die Erforschung psychischen Geschehens und für die Gewinnung psychodynamischer Erkenntnisse dar. Episodische Erzählungen geben Einblick in die innere Welt der Erzähler und ihre Motivationen, Wünsche, Ängste und Konflikte. Die *Erzählanalyse JAKOB* stellt dafür geeignete Prozeduren und Auswertungsmuster zur Verfügung.

Während einige der narrativen Verfahren auch dialogische Gesprächsstellen in die Narrative einbeziehen, beschränkt sich die Erzählanalyse JAKOB auf monologische Erzählungen. Dies ist ein oft kritizierter Punkt, gerade von Gesprächsanalytikern, die den Fokus auf die Interaktion und das Dialogische legen. Was ist die Rolle der *Interaktion* in den Narrativen? Die besondere Form und Funktion der *Erzählung* in der Erzählanalyse JAKOB beinhaltet die Annahme, dass die Erzählerin mit der Ankündigung der Erzählung das Rederecht für die Dauer der Erzählung erhält. Die Bezugnahme auf den Andern ist trotzdem in der Erzählung enthalten; ja die Erzählung

lebt von dieser Bezugnahme, sie richtet sich an das Gegenüber und setzt es als Rezipienten und Mitspieler ein.

Die einzelnen Auswertungsschritte der Erzählanalyse JAKOB ermöglichen teilweise subjektive Beurteilungen, wo objektive Operationalisierungen vorliegen; es besteht deshalb das Bedürfnis, bestimmte Schritte der Analyse reliabler und damit auch objektiver durchführen zu können. Die Kodierung des Erzählvokabulars mit einem vorgegebenen Kategorienschema ist ein solcher Analyseschritt, der objektiviert werden kann, wenn die Zuordnung der Codes zu Ausdrücken festgelegt ist, z. B. in einem Lexikon. Diesen Kodiervorgang sollen das vorliegende Lexikonprojekt und das Lexikon erleichtern und automatisieren helfen. Natürlich gibt es daneben genügend Schritte, vor allem bei der Formulierung der Spielregel und der Wunsch-, Angst- und Konfliktthemen, wo interpretatorischer Spielraum besteht.

Einerseits ist es sicher unumstritten, dass mit einer geeigneten und differenzierten Gestaltung des Lexikons die Zuordnung von Codes zu Text verbessert werden kann, andererseits bestehen grundsätzliche Vorbehalte aus dem Lager der qualitativen Forscher und der Gesprächsanalytiker gegenüber einer a priori vorgenommenen Kodierung. Kritische Stimmen bezweifeln die Plausibilität von a priori im Lexikon festgelegten Kategorien und monieren, dass Wörter und Ausdrücke nur im Kontext und in der Interaktion ihre spezielle Bedeutung erhalten (Auer, 2007; Deppermann, 2006a). Wenn die lexikalische Dimension (Kodierung) ernst genommen werden soll, muss es eine Entwicklung hin zu einem reichhaltigen und überprüften Lexikon geben, das die sonst ausserordentlich aufwendige Auswertung objektiviert und erleichtert. Ein ad hoc-Umgang mit 93 Verbkategorien kann keine brauchbaren Ergebnisse liefern. Weitere grundsätzliche und wichtige Motivationen für die Beschäftigung mit dem Lexikon ergeben sich aus den Überlegungen zum *lexical choice*, die die Vielschichtigkeit der lexikalischen Wahlen aufzeigen und die dahinter stehenden kognitiven Prozesse beleuchten.

2.7 Textanalyse in der Psychotherapie

Während in den bisher dargestellten Theorien und Methoden immer das *Narrativ* oder die *Erzählung* als besondere Diskurs- und Textform im Zentrum stand, geht es im folgenden Kapitel um die Darstellung von *textanalytischen Verfahren*, die grösstenteils mit Computerunterstützung linguistische Merkmale von grösseren Text- oder (transkribierten) Gesprächseinheiten systematisch erfassen und in Bezug auf psychologisch oder psychotherapeutisch relevante Kriterien auswerten (Pennebaker, Mehl & Niederhoffer,

2003; Züll & Mohler, 1992; Züll & Landmann, 2002).

Was in den letzten zwei Dekaden an *textanalytischen Werkzeugen* zur Verwendung mit Computerhilfe entwickelt wurde, wird in einem relativ umfassenden Bericht von Pennebaker et al. (2003) vorgestellt.

Das *Vokabular*, das Menschen in ihrem täglichen Leben verwenden, kann wichtige Aspekte ihrer sozialen und psychischen Welt offenbaren. Mit Computerhilfe können Sprachdaten und linguistische Besonderheiten aus schriftlichen oder transkribierten mündlichen Texten erfasst und ausgewertet werden. Die Autoren fokussieren auf die Art und Weise, wie die vorgestellten Textanalyse-Instrumente den Sprachgebrauch eines Individuums zu Persönlichkeitszügen, sozialen und situationalen Faktoren oder zu psychologischen Interventionen und Therapien in Beziehung setzen. Besonders heben die Autoren um Pennebaker hervor, dass der *Partikelgebrauch* im Gegensatz zu inhaltlich relevanteren Kategorien wie Nomen und Verben psychologisch wertvolle Aufschlüsse vermitteln kann und dass *Partikel als Marker* für Gemütszustände, soziale Identität und kognitive Stile dienen können.

Kritisch beleuchten (Züll & Landmann, 2002, S. 8) die bei Pennebaker et al. vorgestellten Text- und Inhaltsanalysemethoden (cui).²¹ Die meisten Verfahren funktionieren als einwortbasierte, kontextfreie Analysemethoden, die mit grossen Nachteilen behaftet sind. Viel diskutiert wird zum Beispiel immer wieder die Notwendigkeit von Arbeitsschritten wie Lemmatisierung, Disambiguierung, das Erkennen von Negationen und das Auflösen der Referenzen von Pronomina. Diese Probleme sind auch in der Erzählanalyse JAKOB bekannt, und hier bestehen die gleichen Verbesserungspotenziale.

2.7.1 Diktionäre

So genannte *Diktionäre* speichern einen *Spezialwortschatz* für ein gewisses pragmatisch oder domainspezifisch abgegrenztes Gebiet und dienen vor allem der computergestützten Textanalyse (Züll & Landmann, 2002):

Kern des diktionärbasierten Ansatzes ist, wie der Name schon sagt, ein Diktionär (Wörterbuch). Dieses Diktionär enthält Codierregeln und Wortlisten. Mit Hilfe dieser Codierregeln und Wortlisten werden einzelnen Textelementen (z. B. Wörtern oder Phrasen) Codes zugewiesen und die entsprechenden Einheiten damit kategorisiert. (S. 10)

Das *Affektive Diktionär Ulm* (ADU) ist eine Methode der quantitativen Vokabularbestimmung. Mit der *Wortschatzanalyse* wird der manifeste In-

²¹cui – Computerunterstützte Inhaltsanalyse

halt des Textmaterials untersucht. Es werden im Text die einzelnen Wörter gezählt, die mit der Liste im Wörterbuch übereinstimmen. Der Kontext wird nicht beachtet (Hölzer, Scheytt & Kächele, 1992). Die Begriffsliste des ADU besteht aus «all [den Wörtern] eines Sprechers, die auf der Einzelwortebene eine emotionale Konnotation besitzen: als die Summe aller verwendeten Gefühlsworte» (S. 133). Aufgrund der Beschränkung auf das affektive Vokabular werden lediglich ca. 4% der Wörter eines Textes untersucht, in Therapiegesprächen macht der Anteil etwas mehr aus, ca. 10%. Grundlage des ADU bildet ein Klassifikationsschema mit acht *Grundemotionen* (Klassifikation nach Dahl, 1978): *Liebe, Begeisterung, Wut, Furcht, Freude, Zufriedenheit, Angst, Depression*. Diese Grundemotionen bestehen aus *Objektemotionen* (it-emotions, 1-4) und *Selbstemotionen* (me-emotions, 5-8). Durch Rater wurden über 2000 *Emotionswörter* aus unserer Sprache den acht Grundemotionen zugeordnet. Die Kodierung von Texten ist äusserst aufwendig, weil einerseits viele Emotionswörter nicht erkannt werden (wie z. B. metaphorische oder idiomatische Ausdrücke), andererseits gibt es viele falsche Kodierungen, weil das System die Wortformen mangels Lemmatisierungsprozeduren nicht identifizieren kann (Bsp. «Macht» vs. «macht») (Wipf, 2010). Die Kategorien und Worteinträge des ADU sind nicht öffentlich dokumentiert.

Eine neuere Studie beschreibt die Anwendung des ADU für eine textanalytische Untersuchung zum Thema «Emotionen im Lebensrückblick alter Menschen» (Wipf, 2010):

Die Ergebnisse weisen auf *Emotionsveränderungen* über die verschiedenen Lebensabschnitte hin. Die Erzähler erzielen im Alter Höchstwerte für *Zufriedenheit* und die positiven Me-Emotionen. Bei der Emotion *Wut/Zorn* liegen die Höchstwerte im Lebensabschnitt Kindheit und die niedrigsten Werte im Alter. Dabei ist ein stetiger Rückgang der Werte von der Kindheit bis ins hohe Alter zu beobachten. Beim Vergleich der Kategorien *Ich* (Subjekt) versus *Andere* (Objekte) konnten ebenfalls Unterschiede gefunden werden. So verzeichnet die Kategorie *Ich* höhere Werte für *Freude, Ängstlichkeit* und die positiven Me-Emotionen. Für die Kategorie *Andere* wurden höhere Werte für *Liebe, Begeisterung* und die positiven It-Emotionen gefunden. Die Ergebnisse bestätigen die Hypothese, dass im Alter weniger *Ärger* (*Wut/Zorn*) berichtet wird.

Auf einer ähnlichen Grundlage arbeitet das *Therapeutic Cycles Model* (TCM) (Mergenthaler, 1996). Emotions- und Abstraktionsmuster (mehr-

maliges Vorkommen von Emotions- oder Abstraktionswörtern in einem Gesprächsabschnitt) werden auf dem Hintergrund eines entsprechenden Diktionärs für das Emotionsvokabular (siehe oben) und für das Abstraktionsvokabular in Therapiegesprächen identifiziert. Das Modell nimmt an, dass sich im Therapiegespräch Phasen mit einem hohen Anteil an Emotionsausdrücken (*experiencing*) mit Phasen mit einem hohen Anteil an Abstraktionsausdrücken (*reflecting*) abwechseln und dass daraus Schlüsse über den Therapieverlauf gezogen werden können (Fontao & Mergenthaler, 2008; Lepper & Mergenthaler, 2007).

Auch die *Gottschalk-Gleser-Angstskalen* und das *Dresdener Angstwörterbuch* basieren auf Diktionären (Berth, Balck & Brähler, 2005). Das Dresdener Angstwörterbuch (DAW) ist die computergestützte deutsche Version des bekannten Gottschalk-Gleser-Verfahrens. In der deutschen Adaptation wurden vor allem die Affekte *Angst* und *Aggressivität* übernommen. Das Wörterbuch enthält 4070 Wortstämme, die den einzelnen Angstskalen Todesangst, Verletzungsangst, Trennungsangst, Schuldangst, Angst vor Scham und diffuse Angst zugeordnet werden können. Weiter wurde, wie in den Originalskalen, berücksichtigt, auf wen sich die Angst bezieht und wie stark diese ausgeprägt ist. Die Validität und Reliabilität des DAW wurden überprüft und als gut bewertet (S. 494). Das Gottschalk-Gleser-Verfahren wurde in zahlreichen Studien angewendet.

Linguistic Inquiry and Word Count (LIWC)

Der Psychologe James Pennebaker hat mit seinen Mitarbeitern über längere Zeit das Textanalyseprogramm *Linguistic Inquiry and Word Count* (LIWC) entwickelt, das auf einem vorgegebenen Diktionär beruht und elektronische Texte auf Einwortbasis analysiert und auswertet.

Die täglichen Einträge von 1084 Usern eines Online Blogs²² wurden für eine Zeitspanne von 4 Monaten mit LIWC ausgewertet, davon 2 Monate vor und 2 Monate nach den Terrorattacken des 11. September 2001. Linguistische Analysen der Blogeinträge zeigten ausgeprägte sprachliche und emotionale Veränderungen als Antwort auf die Attacken. Kurzfristig äusserten die Schreiber mehr negative Emotionen, zeigten sich kognitiv und sozial engagierter und schrieben mit mehr psychologischer Distanz. Nach 2 Wochen kehrten die Emotionen und die soziale Engagiertheit zur Grundlinie zurück, und der Gebrauch eines kognitiven und

²²Users of livejournal.com

analytischen Vokabulars fiel sogar unter die Grundlinie (Cohn, Mehl & Pennebaker, 2004).

Die Autoren konzentrierten sich für ihre Analyse auf vier in Sozial- und Klinischer Psychologie anerkannte thematische *Coping-Konzepte*: Emotional processing, cognitive processing, social and interpersonal concerns, psychological closeness or distance (Cohn et al., 2004, S. 688). *Emotional processing* bezeichnet eine temporäre depressionsähnliche Verstimmung nach einem traumatisierenden Ereignis. *Cognitive processing*: Die kognitive Aktivität nimmt nach einem traumatischen Ereignis eher zu, weil die Betroffenen das Vorgefallene zu verstehen versuchen und einen Sinn in die Ereignisse legen, indem sie z. B. ein kohärentes Narrativ entwickeln (developing a narrative). Traumatische Ereignisse können die soziale Interaktion fördern, weil die Menschen Unterstützung suchen, indem sie die Ereignisse mitteilen und kollektiv verarbeiten. Schliesslich ist eine verbreitete Abwehrstrategie, sich von den Ereignissen zu distanzieren. Dies kann durch die linguistische Grösse *verbal immediacy* vs. *verbal nonimmediacy* geprüft werden, was entsprechend mehr oder weniger psychologischer Distanzierung entspricht (high/low psychological distance).

Tabelle (2.2) zeigt einen Ausschnitt der Kategorien und Beispiele der Gruppe *Psychological Processes* des englischen Diktionärs von LIWC.²³ Es existiert ebenfalls eine *deutsche Version* des LIWC, d. h. eine Version mit einem deutschen Wörterbuch, die der oft verwendeten und geprüften englischen Version ebenbürtig sein soll (Wolf et al., 2008).

Tabelle 2.2: Beispielkategorien des LIWC: Psychological Processes

Category	Examples	Words in cat.	Validity
Social processes	mate, talk, they, child	455	
Family	daughter, husband, aunt	64	.87
Humans	adult, baby, boy	61	
Affective processes	happy, cried, abandon	915	
Positive emotion	love, nice, sweet	406	.41
Negative emotion	hurt, ugly, nasty	499	.31
Anxiety	worried, fearful, nervous	91	.38
Anger	hat, kill, annoyed	184	.22
Sadness	crying, grief, sad	101	.07
Cognitive processes	cause, know, ought	730	
Insight	think, know, consider	195	
Causation	because, effect, hence	108	.44

²³ *Words in cat.* = Anzahl Wörter in dieser Kategorie; *Validity* = Validitätsmass auf Grund von Raterurteilen.

2.8 Fazit zu diktionärbasierten Verfahren

Diktionärbasierte Verfahren sind fehleranfällig, weil sie den Satz-, Situations- und Diskurskontext nicht berücksichtigen; dies wird in der zitierten Studie von Wipf (2010) anschaulich demonstriert. Die Limitierung erwähnen auch Rullkoetter et al. (2008, S. 285). Die Autorinnen monieren in ihrer textanalytischen Studie, dass wohl Wörter gezählt werden können, dass aber das Instrument Fehler macht bei der Beurteilung von mehrdeutigen Wörtern oder Sätzen, wenn die gefundene Wortbedeutung im Kontext nicht zutrifft, oder wenn der Satzsinne durch spezielle Intonation vom gefundenen abweicht (siehe auch Züll & Landmann, 2002, S. 8).

Die oben erwähnten Publikationen zum LIWC und zum ADU können zwar auf interessante Ergebnisse verweisen, die linguistische Evidenz der Befunde ist jedoch klein. Die erwähnten Instrumente eignen sich nicht für linguistische Analysen, die mehr als Hinweise für ganz spezielle und auf ein enges Thema beschränkte Phänomene liefern sollen.

Für die *Erzählanalyse JAKOB* ergibt sich daraus eine starke Motivation, das Lexikon so weiter zu entwickeln, dass möglichst viele dieser Unzulänglichkeiten verbessert werden können und dass durch eine geeignete Gestaltung der Lexikoneinträge differenzierende Merkmale geschaffen werden können, die eine verbesserte Zuordnung von Textinstanzen und Lexikoneinträgen erlauben.

3 Linguistik und Gesprächsanalyse

Das folgende Kapitel präsentiert Studien und Methoden aus den Gebieten der *Gesprächsforschung* und *Interaktionalen Linguistik*. Diese gesprächslinguistischen Disziplinen untersuchen Gesprächsdaten, um «dialogisches sprachliches Handeln in sozialen Situationen systematisch zu beschreiben und zu erklären» (Brinker & Sager, 2006), also mit einem etwas anderen Fokus als die im vorherigen Kapitel beschriebenen Erzähl- und Textanalysen, die sprachliche Produktionen zum Teil losgelöst vom dialogischen Gesprächsablauf betrachten. Gesprächsforschung umfasst im Folgenden sowohl die *Konversationsanalyse* (conversation analysis), wie sie im Rahmen der Ethnomethodologie im englischen Sprachraum entwickelt wurde, als auch den im deutschen Sprachraum gebräuchlichen Begriff der *Gesprächsanalyse*. In der Interaktionalen Linguistik werden ebenso Gespräche untersucht; die methodische Verankerung liegt aber hier mehr auf linguistischen als auf ethnologischen Konzepten. Ebenfalls unter dem Thema Gesprächsforschung wird die *Gesprochene-Sprache-Forschung* subsumiert, die zwar schon ab den 1960er Jahren entstanden ist, in den letzten zehn Jahren aber im Rahmen der Gesprächsforschung eine Aktualisierung erlebt (Deppermann, 2007; Günthner, 2007a; Schwitalla, 2001). Das Phänomen der *Bedeutungskonstitution*, d. h. die Frage, wie Bedeutung im Gespräch entsteht, ist eine Grundfrage aller gesprächslinguistischen Ansätze, erhält in der Arbeit am Lexikon eine prominente Bedeutung und wird deshalb ausführlich dargestellt. Nicht nur das *Wie* der Bedeutungsentstehung, sondern auch das *Wann* ist eine wichtige Frage, über die keineswegs Konsens besteht und die in der *Emergent Grammar* thematisiert wird. Während sich das Interesse der traditionellen Gesprächsforschung zu einem grossen Teil auf lokale Eigenschaften der *Prosodie*, der *Syntax* und der *Semantik* bezieht, besteht das Bedürfnis und die Notwendigkeit, den Fokus vom Wort in Richtung Satz und *Diskurs* zu erweitern und satzsemantische und pragmatische Eigenschaften und Merkmale der *Gesprächssituation* in die Analyse einzubeziehen, was durch die in den letzten Jahren gestiegene Verfügbarkeit von elektronischen Gesprächskorpora begünstigt wird.

Obschon *gesprochene Sprache* in der menschlichen Entwicklung über Zehntausende von Jahren der elementare Sprachausdruck war, beschäftigt sich die traditionelle westliche Linguistik seit Antike und Mittelalter vorwiegend mit der Schriftsprache, und das Primat der Schriftsprache beeinflusst tiefgreifend unser Denken über die Sprache (Tomasello, 2003). Erstens war aber das Schreiben historisch gesehen nötig, um Aspekte der gesprochenen Sprache überhaupt ins Bewusstsein zu bringen, zweitens kann keine Schriftsprache *alle* Nuancen der gesprochenen Sprache abbilden, und drittens werden die durch Schriftsprache abgebildeten Eigenschaften der gesprochenen Sprache irrtümlich als komplettes Sprachmodell angesehen (Olson, 1994, zit. bei Tomasello, 2003, S. 3). Tomasello charakterisiert die *gesprochene Sprache* (SSS, spontaneous spoken speech) mit den folgenden psycholinguistisch orientierten Eigenschaften:

- Gesprochene Sprache bildet keine Sätze, sondern äussert sich in *intonation units*; diese Einheiten sind prosodisch und semantisch kohärent und enthalten in der Regel nur *eine* neue Information (vgl. Du Bois, 2003)
- Gesprochene Sprache besteht selten aus prototypischen vollständigen Sätzen; meistens wird ein neuer *Referent* in einer Intonationseinheit eingeführt, in der nächsten dann etwas über diesen ausgesagt (Prädikat).
- Gesprochene Sprache verwendet seltener *Relativsätze* und dies nur dann, wenn der Referent schon eingeführt ist.
- Gesprochene Sprache verwendet wenig prototypische *transitive Sätze*, weil tendenziell nicht zwei neue Referenten gleichzeitig eingeführt werden (siehe oben).
- Gesprochene Sprache und Schriftsprache verwenden nicht die gleiche *Grammatik*. (S. 3)

Die im Folgenden vorgestellten linguistischen Disziplinen der Interaktionalen Linguistik und Gesprächsanalyse befassen sich hauptsächlich mit gesprochener Sprache und bestätigen die von Tomasello postulierten Unterschiede teilweise, auch wenn diese vielleicht vom linguistischen Standpunkt aus gesehen etwas übers Ziel hinausschiessen.

3.1 Interaktionale Linguistik

Die *Interaktionale Linguistik* (IAL), programmatisch dargestellt z.B. bei Selting und Couper-Kuhlen (2000), befasst sich mit der «Struktur und Verwendung der Sprache in der sozialen Interaktion» und umfasst sämtliche linguistischen Aspekte wie Prosodie, Syntax, Phonetik, Phonologie, Morphologie, Lexik, Semantik und Pragmatik. Es geht zentral um die Kontextualisierung von Sprache in der Situation und Interaktion, im Gegensatz zur Betrachtung von Sprache als von der Verwendung unabhängiges kognitives System (Sprachkompetenz vs. Performanz). Im Unterschied zur Konversationsanalyse, die sich mit interaktiven Mechanismen und der Organisation des Gesprächs beschäftigt, fokussiert die IAL auf die sprachliche Realisierung. In Europa als *interactional linguistics* bezeichnet, in Amerika als *discourse-functional linguistics* (Ford, Fox & Thompson, 2002; Hakulinen & Selting, 2005a), ist die IAL entstanden aus dem Bedürfnis, den ethnomethodologischen und konversationsanalytischen Sichtweisen und Methoden zu mehr Bedeutung in der Linguistik zu verhelfen und Perspektiven auf die Struktur und Verwendung von Sprache in ihrer natürlichen Umgebung, der sozialen Interaktion, hinzuzufügen.¹

Schwerpunkte der IAL bilden einerseits die Untersuchung der interaktiellen oder konversationellen Aufgaben in Gesprächen (Welche sprachlichen Mittel werden für Umsetzung verwendet?), andererseits die Betrachtung der linguistischen Elemente oder Konstruktionen und der Rollen, die diese Phänomene in der Interaktion spielen (Selting & Couper-Kuhlen, 2000, 2001). Die Interaktionale Linguistik untersucht

- i) what linguistic resources are used to articulate particular conversational structures and fulfil interactional functions?
- ii) what interactional function or conversational structure is furthered by particular linguistic forms and ways of using them? (Couper-Kuhlen & Selting, 2001)

Als wichtige Grundannahmen gelten gemäss Selting und Couper-Kuhlen (2000, S. 81) für die IAL:

- Sprachstrukturen sind *interaktionale*, nicht nur individuelle Produkte/Leistungen.

¹Stichworte: Prosody in conversation, Interaction and Grammar, Studies in Discourse and Grammar.

- Sprachstrukturen sind flexibel und anpassungsfähig.
- Es besteht ein wechselseitiger Zusammenhang zwischen *Sprachstrukturen* und dem sequentiellen *Interaktionskontext*. Sprachstrukturen sind auf die Organisation der Interaktion zugeschnitten und die Organisation der Interaktion wird durch Sprachstrukturen in bestimmten sequentiellen Kontexten hergestellt.
- Die *Bedeutungskonstitution* und die Interpretation von Aktivitäten oder Handlungen sind Ergebnisse des Einsatzes von linguistischen Verfahren und Praktiken in bestimmten Kontexten.

Sprachstrukturen sind gemäss IAL *emergent*, d. h. die linguistischen Kategorien und Einheiten werden mindestens teilweise in der Interaktion hergestellt. Es stellt sich deshalb die Frage nach der Relevanz kognitiver Konzepte und deren Zusammenhang mit interaktionalem Handeln. Wie soll das Verhältnis zwischen *Interaktion* und *Kognition* bestimmt werden, wie kann die lokale, situative *Emergenz* sprachlicher Strukturen mit kognitiven Theorien, z. B. zur schemaorientierten Produktion und Planung sprachlicher Strukturen und Handlungen, in Einklang gebracht werden? Verschiedene Studien untersuchen eben diesen Zusammenhang zwischen linguistischen (grammatischen) Strukturen und Gesprächsprozessen (Deppermann, Fiehler & Spranz-Fogasy, 2006). Deppermann (2007, S.20) präsentiert eine Gegenüberstellung von IAL und kognitiven Ansätzen und arbeitet die Gemeinsamkeiten und Unterschiede heraus. Die IAL ist ein funktionalistischer Ansatz und lehnt formale Modelle im generativistischen Sinne ab. Mehr Berührungspunkte gibt es mit Ansätzen der *kognitiven Linguistik*, wie der *Konstruktionsgrammatik* oder der Kognitiven Grammatik. Die Gemeinsamkeiten bestehen darin, dass Sprache primär funktional, als Mittel der Kommunikation angesehen wird (S. 20). Die Grundannahmen der Interaktionalen Linguistik lassen sich sehr gut mit Grundsätzen und Ideen der Konstruktionsgrammatik kombinieren; interaktionale und kognitive Aspekte der sprachlichen Aktivitäten können integriert dargestellt werden (vgl. Kapitel 6.5), weil kognitive Prozesse Voraussetzungen für interaktionale Gesprächsprozesse sind (Deppermann, 2006a).

Aijmer und Stenström (2005, S. 1745) stellen die IAL als Theorie und Methode in den Kontext von Soziolinguistik, Diskursanalyse, Konversationsanalyse und allgemein der Erforschung von gesprochener Sprache:

Interactional linguistics [...] combines a sociolinguistic and a linguistic approach, aiming at a grammatical description of <talk-

in-interaction>. As in CA, the focus is on the analysis of naturally occurring spoken data, now with particular attention paid to the way in which syntax, lexis, and prosody are used and their role in the conversational organization.

3.2 Gesprächslinguistik, Gesprächsforschung

Unter dem Begriff *Gesprächsforschung*² hat sich in den letzten 30 Jahren eine Vielzahl von theoretischen Konzepten und praktischen Analysemethoden entwickelt. Die *Gesprächslinguistik* betont ähnlich wie die Interaktionale Linguistik die *Gebrauchsbasiertheit* sprachlicher Strukturen gegenüber der Sprache als System, stellt der grammatischen Kompetenz des Systems das sprachliche Handeln und das Kommunikationsereignis gegenüber (Deppermann & Spranz-Fogasy, 2001, S. 1148). Auch die *Gesprächsanalyse* (GA) sollte stets als *funktionale Analyse* durchgeführt werden, mit engem Bezug zum Sprachgebrauch und den kommunikativen Funktionen, nicht als Untersuchung eines autonomen Systems (Imo, 2007, S. 15). Die Gesprächsanalyse ist aus der *Konversationsanalyse* (CA) entstanden, kann aber nicht damit gleichgesetzt werden. Die Hintergrundannahmen und die Methoden der Gesprächsanalyse haben sich vor allem in der deutschsprachigen Forschung etabliert. Ein geschichtlicher Überblick und eine ausführliche Gegenüberstellung von Konversationsanalyse und Gesprächsanalyse findet sich bei Deppermann (2006c). Brinker und Sager (2006) stellen die *linguistische Gesprächsanalyse* als Forschungsrichtung in einer Einführung und einem forschungsgeschichtlichen Überblick dar.

Georgaca und Avdi (2009) formulieren in einer Übersichtsstudie zu sprachbasierten Analysemethoden in der Psychotherapieforschung drei Schwerpunkte, charakterisieren damit Erzählanalyse, Diskursanalyse und Konversationsanalyse und schlagen folgende Abgrenzung der Begriffe vor:

1. In der *Erzählanalyse* werden Therapiegespräche inhaltlich untersucht, mit dem Ziel, die Probleme der Klienten zu benennen und zu formulieren.
2. In der *Diskursanalyse* wird der Therapieprozess untersucht: Wie geschieht Veränderung im Laufe der Zeit?

²Gülich (2001, S. 1086) betont die Verwandtschaft der Begriffe Gesprächslinguistik, Gesprächsanalyse und Gesprächsforschung und der damit beschriebenen Phänomene.

3. Die *Konversationsanalyse* konzentriert sich auf einzelne Redebeiträge von Patient und Therapeut, auf die Interventionen des Therapeuten und die Wirkung auf den Therapieprozess (Georgaca & Avdi, 2009, S. 233).

Deppermann (2006c) ordnet die verschiedenen methodischen Ansätze aus gesprächsanalytischen Studien zu einer expliziten Methodensammlung zur Gesprächsanalyse, zu einer Art «Werkzeugkasten». Die Gesprächsanalyse untersucht grundsätzlich, wie Menschen Kommunikation gestalten und dadurch Wirklichkeit herstellen. Kallmeyer (1985; zit. nach Deppermann, 2006c, S. 9) unterscheidet sechs Ebenen der Herstellung von sprachlicher Interaktion, die als *Gesprächspraktiken* bezeichnet werden. Für das vorliegende Projekt eines Konstruktionslexikons bilden diese Ebenen wichtige Fragen an das Gesprächsmaterial für die Charakterisierung von Konstruktionen (die Ausdrücke in Klammern beziehen sich auf Lexikoneigenschaften, die später in der praktischen Implementierung verwendet werden, siehe Kap. 9).

1. *Gesprächsorganisation*: formale Abwicklung, Eröffnung, Beendigung des Gesprächs.
2. *Darstellung von Sachverhalten*: Gesprächsthemen und -inhalte (<topic>), Gesprächsform: Argumentieren, Beschreiben, Erzählen (<textType>).
3. *Handeln*: Ziel und Zweck des Gesprächs (<domain>): Arztgespräch, Streitschlichtung etc.
4. *Soziale Beziehungen* zwischen den Gesprächsbeteiligten, z. B. Macht Vertrauen, Sympathie (soziale Rollen), und Identität derselben.
5. *Modalität* des Gesprächs: Ernst, Spass, Spiel, emotionale und stilistische Beteiligung, Betroffenheit, Ärger, Zurückhaltung (<register>, <textType>).
6. *Herstellung von Reziprozität*: Verständigung und Kooperation.

Wesentliches Kennzeichen der Gesprächsanalyse ist der Fokus auf die sequenzielle und dialogische Einbettung der untersuchten Gesprächsausschnitte sowie auf ihre kontextspezifischen und interaktionalen Eigenschaften. Neben den Gesprächspraktiken sind auch *Situation* und *Situiertheit* wichtige Merkmale der Gesprächsanalyse (Deppermann & Spranz-Fogasy, 2001, S. 1148). Die Themen (und damit die *Bedeutung* von Äusserungen und

Ausdrücken) verändern sich im Gespräch fortwährend, Referenzobjekte und Sinngehalt müssen von den Gesprächsbeteiligten laufend neu bestimmt und analysiert werden. Imo (2007, S. 12–13) postuliert eine *interaktionale Semantik*: Der semantische Gehalt von Wörtern und Phrasen ist stark von der Situation abhängig und wird teilweise erst in der Interaktion festgelegt (vgl. Deppermann, 2006d). Andererseits kann Bedeutung nicht ausschliesslich ad hoc und lokal entstehen, da die Kommunikation dadurch sehr umständlich und langsam würde. Im allgemeinen wird eine *Kernbedeutung* postuliert (basale Kernsemantik), die im Rahmen der Situation von Sprecher und Hörer angepasst und verändert werden kann. Die Kernbedeutung ist eher vage und entspricht einem Lexikoneintrag, die *lokale Semantik* ergänzt, verändert oder ersetzt die Kernbedeutung dynamisch im Kontext der Situation (Deppermann, 2006d; Imo, 2007).

Die *Kontextabhängigkeit* der Bedeutung von sprachlichen Äusserungen ist nicht nur für die Gesprächsanalyse ein wichtiges (Postulat), auch in der *Konstruktionsgrammatik* ist der Kontext wichtig. Gesprächsanalyse und Konstruktionsgrammatik ergänzen einander theoretisch und methodisch und können gewinnbringend kombiniert werden (Deppermann, 2006a; Imo, 2007). Sprachliche Äusserungen können nicht ohne ihren Kontext betrachtet werden; darunter fällt sowohl der Kontext der Situation als auch der weitere Kontext der Kultur (Linell, 1998, 2009). Die akribische gesprächsanalytische Untersuchung einzelner Textstellen muss ergänzt werden mit Informationen und Wissen über den engeren und weiteren Kontext, kommunikative Praktiken und Gattungen oder Genres. Diese erweiterte Betrachtungsweise wird auch in der Konstruktionsgrammatik angestrebt (Imo, 2007; Östman, 2005, S. 14). Die Form/Funktionsbeziehungen gelten über die Konstruktion und den Satz hinaus auch für den Diskurs (vgl. Kap. 6.7).

Eine *Grammatiktheorie der Gesprochenen Sprache* müsste den dynamischen Charakter des «talk-in-interaction» in die Konstruktionsbeschreibung aufnehmen (Imo, 2007, S. 12–13). *Gesprochene Sprache* wird in Echtzeit produziert und rezipiert und ist deshalb von der dialogischen Situation bestimmt. Die Kommunikation erfordert schnelle Reaktionen, sprachliche Strukturen müssen sofort abgerufen und geeignet kombiniert werden können. Auer (2007, S. 96) spricht von *On line-Syntax*. Im Gesprächsverlauf werden von den Sprechenden mögliche oder beabsichtigte Gesprächswendungen durch *Projektionen* eingeleitet. Eine syntaktische Gestalt wird eröffnet (z. B. über die Valenz). Die Gesprächspartner orientieren sich an diesen verfestigten Schemata, die selber wiederum aufgrund ihrer gestalthaften Qualität Projektionen über den weiteren Interaktionsverlauf erlauben (siehe Kapitel 3.2.1).

Dem Umstand, dass die mündliche Sprachproduktion ihre Produktionsregeln anpassen und verändern kann, wird in der *Emergent Grammar* Rechnung getragen (Hopper, 1987): Grammatik wird in der Interaktion immer wieder neu hergestellt. Wenn bestimmte Sprachmuster häufig wiederholt werden, bilden sich entsprechende Grammatikregeln; wenn die Sprachmuster umgekehrt nicht gebraucht werden, verschwinden sie mit der Zeit aus der Grammatik (vgl. Kap. 6.8). Grammatik ist in diesem Sinne das Inventar der aktuell verwendeten sprachlichen Routinen (Imo, 2007, S. 17–18). Im gleichen Sinn argumentiert Hoey (2005) mit seiner Theorie des *lexical priming*:

Every word is primed for use in discourse as a result of the cumulative effectives of an individual's encounters with the word. Primings can nest, in that the resulting word sequences from initial primings are themselves the object of further priming. (Hoey, 2005, S. 13)

3.2.1 Konversationsanalyse, Conversation Analysis

Obschon die Konversationsanalyse historisch gesehen der Gesprächsanalyse vorausgeht, stelle ich sie im Folgenden unter der Gesprächsanalyse vor. Zentrales Thema der *Konversationsanalyse* (*conversation analysis*, CA) ist das Gespräch als Interaktion, «the Study of Talk-in-Interaction» (ten Have, 2007). Der Begriff Konversationsanalyse umfasst in einem weiteren Sinne jede Untersuchung von Phänomenen, die entstehen, wenn Leute zusammen sprechen («any study of people talking together»); in einem engeren Sinne ist damit die CA als Forschungsrichtung gemeint, wie sie sich ab 1960 in Kalifornien mit den Hauptakteuren Sacks, Schegloff, Jefferson und Goffman entwickelt hat (ten Have, 2007, S. 3–5). Von Anfang an gab es zwei Typen von CA: auf der einen Seite die CA als «pure» Untersuchung der Interaktion an sich, auf der anderen Seite die angewandte Konversationsanalyse («applied CA»), die als Forschungsinstrument die Funktion von Gesprächsphänomenen in sozialen Zusammenhängen und in der Interaktion untersucht («management of social institutions in interaction») (ten Have, 2007, S. 6–7).

Wichtige Grundsätze der CA auf dem Hintergrund von Soziologie und Linguistik sind: CA soll nahe an den untersuchten Phänomenen bleiben und verwendet deshalb reale Gesprächsdaten. Der Fokus der CA liegt auf der Interaktion und auf den kollektiven Abläufen und Prozessen, die diese bewirken (interaktive Aushandlung, lokales Management). CA will nicht er-

klären, warum Leute so und so handeln, sondern beschreiben, wie sie es tun. CA untersucht die Sprache in der Interaktion mit dem Anliegen, die Interaktanten zu verstehen (ten Have, 2007, S. 9). Transkripte sind beispielhafte Einzelfälle (specimen), nicht Ausdruck von verborgenen Zuständen oder Reflexionen, und auch keine Instanzen von Klassen. CA zeigt Kollektionen von sozialen Praktiken, Regularitäten und normativer Handlungsorganisation auf (ten Have, 2007, S. 35–36). Ein wichtiger Begriff für die CA ist die *Projizierbarkeit*: Das Format einer Konversationseinheit erlaubt Voraussagen über ihr mögliches Ende (Projektion) und mögliche Fortsetzungen (Fischer, 2007).

Einen wichtigen Einfluss auf die CA hatte die *Ethnomethodologie* (Harold Garfinkel), mit ihrer Methode der Beobachtung und der *Kategorienanalyse* (*membership categorization analysis*, *MCA*). Kategorien müssen in jeder Situation neu entwickelt und nur aus dem Datenmaterial erschlossen werden. Deshalb ist CA eine Mischung aus Empirie und Phänomenologie; je wissenschaftlicher der Anspruch, desto empirischer ist meistens das Vorgehen. Auf diesem Hintergrund ist auch die Abwehrhaltung vieler Konversationsanalytiker und Ethnomethodologen gegen eine rein methodische Anwendung der CA («mechanical use of CA») zu verstehen (die allgegenwärtige Angst vor der Kodierung) (ten Have, 2007, S. 48–49).

Aus Sicht des Psychotherapieforschers beleuchtet Streeck (2008, S. 174–176) das konversationsanalytische Vorgehen, welches ein interaktionales Modell der Psychotherapie favorisiert. Alle Ereignisse in der Therapie sind Ko-produktionen der Beteiligten (two-body psychology). Therapeutische Veränderungen werden nicht durch Narrative und Deutungen bewirkt, sondern durch implizites Beziehungswissen aus der therapeutischen Situation und aus Situationen mit andern. Ausserdem stellt Streeck fest, dass konversationsanalytisch in der Regel kleine, lokale Gesprächsausschnitte untersucht werden, und er moniert, dass CA auch grössere Einheiten der therapeutischen Kommunikation ins Auge fassen sollte (Streeck, 2008, S. 183–184).

Methodologisches zur Konversationsanalyse

Für die an der *Ethnomethodologie* orientierte Konversationsanalyse wird auch der Begriff *Alltagsmethodologie* verwendet (Gülich, 2001). Dieser Begriff bezeichnet «die von den Mitgliedern einer Gesellschaft im Handlungsvollzug praktizierte Methodologie...» (Bergmann 1981, zit. nach Gülich, 2001, S. 1086), also ein Phänomen, das sich auf das Handeln der Individuen im Alltag bezieht, und nicht so sehr auf eine wissenschaftliche Methode. Der Begriff Alltagsmethodologie beinhaltet, dass die Aktivitäten der Gesprächs-

teilnehmer zugleich erklärbar und darstellbar gemacht werden (accountable).³ Da die Konversationsanalyse in der Ethnomethodologie begründet ist, erschliesst sich die Methodologie letztlich immer erst in der Forschungspraxis. Allerdings gibt es wenig programmatische Äusserungen dazu, wie man sich dies im Einzelnen vorzustellen hat. Die Analyse-Methoden unterstehen dem *unique adequacy requirement*, d. h. sie müssen dem Gegenstand angemessen sein. Nach Gülich (2001, S. 1088) sind die wichtigsten methodischen Prinzipien der CA:

- empirischer Charakter der Untersuchung
- natürliche Gespräche
- sequenzielle Analyse
- Berücksichtigung der interaktiven Dimension
- Aufdeckung der Ordnungsstrukturen
- Entwicklung der Fragestellungen aus den Daten selbst
- Rekurs auf die Kategorien der Teilnehmer

Der Rekurs auf die *Teilnehmerkategorien* (Sprecherkategorien) wird auch von andern Autoren als wichtiges methodisches Prinzip der CA bezeichnet (Fischer, 2007), d. h. die Kategorien und Einheiten, an denen sich die Gesprächsteilnehmer orientieren, sollen nicht ausserhalb des Kontextes erfunden werden.

Peräkylä (2004) untersucht mit konversationsanalytischen Methoden interaktionale Aspekte der psychoanalytischen Deutung. Mit Hilfe der CA zeigt er die interaktionalen Praktiken auf, mit denen die Verlinkung bewerkstelligt wird, so dass es dem Patienten möglich wird, vergangene und aktuelle Erfahrungen zu verbinden. CA kann allerdings nicht messen (und beschäftigt sich auch nicht damit), ob die Deutungen einen positiven therapeutischen Effekt haben, sondern beschreibt diese ausschliesslich.

Trotz oder gerade wegen der grossen Verbreitung konversationsanalytischer Ansätze und Hintergründe in Theorien und Methoden der Gesprächsanalyse mangelt es nicht an Kritik an der Konversationsanalyse. So äussern Aijmer und Stenström (2005, S. 1744):

³ *account* und *accountable* sind zentrale Begriffe für die CA (Gülich, 2001, S. 1087).

Despite its many advantages, CA has been criticized for its lack of <systematic analytical categories>, its <fragmentary focus> and its <mechanistic interpretation of conversation>, which precludes a comprehensive quantitative analysis.

Der deskriptive Charakter vieler konversationsanalytischer Studien ist ebenfalls ein häufig geäußelter Kritikpunkt. Die Aufgabe der CA ist jedoch gerade die genaue Beschreibung der Praktiken, durch welche die Gesprächsteilnehmer ihre Aufgaben erfüllen (Peräkylä, 2004, S. 303).

Auf konversationsanalytischem Hintergrund diskutieren Kurri und Wahlström (2007) die «Therapie als Konversation», die es explizit ablehnt, das Gespräch nach mentalen Strukturen und Defizienzen des Klienten zu hinterfragen⁴ (im Gegensatz zu erzählanalytischen Ansätzen). Psychotherapie wird als Heilmittel angesehen in Situationen, in denen die Klienten über eine verminderte *agency* in ihrem Leben verfügen. Die traditionelle Sicht vermutet verminderte *agency* auf Grund von Konflikten und Defizienzen in der kognitiven und affektiven Struktur der Persönlichkeit. Dagegen argumentieren die Autoren, dass die Klienten persönliche Verantwortung übernehmen möchten, indem sie ihre Handlungen als intentional und gut überlegt darstellen. Die Konstruktion von aktiver *agency* im Gespräch kann als Vorstufe zur realen Übernahme von Verantwortung angesehen werden. Thema der Studie ist *agentless talk* (unpersönliche Konstruktionen wie Passiv, Nominalisierung, Verwendung von Indefinitpronomen) als Extrembeispiel für die Abwesenheit von *agency*.

Konversationsanalyse und allgemeine Linguistik

Gegenüber der (klassischen) Linguistik haben viele Vertreter der CA Vorbehalte, während sich andere vermehrt der Linguistik zugewandt und die Ideen der CA z. B. mit *funktionalen Linguistiktheorien* verbunden haben (Ford & Thompson, 1996; Ford et al., 2002; Ford, Fox & Thompson, 2003; Ochs, Schegloff & Thompson, 1996). Diese Autorinnen haben die CA mit Gewinn für beide Seiten näher zur Linguistik gebracht.

Grammatik und Interaktion beeinflussen sich gegenseitig (Ochs et al., 1996; Schegloff, Ochs & Thompson, 1996), die soziale Interaktion wird als Basis für Kreativität, Lernen und Veränderung in der Sprache begriffen.

⁴ «This postmodern or linguistic turn in therapy research (Kvale 1992; Rorty 1979) rejects the quest to go beyond talk into the inner mental structures of the client. What the client says is not taken to reflect inner psychic realities, but is considered intelligible as a part of the conversation» (Kurri & Wahlström, 2007, S. 315).

Ford und Thompson (1996) beschreiben den Sprecherwechsel und das Sprechermanagement (turn organization) auf dem Hintergrund von syntaktischen, intonationalen und pragmatischen Ressourcen.

Die Möglichkeiten (und das Interesse) der Konversationsanalyse, mit verschiedenen Richtungen der *Linguistik* zu kommunizieren, diskutiert der Sammelband von Hakulinen und Selting (2005b). In ihrer Einführung stellen Hakulinen und Selting (2005a) die CA in den Kontext verschiedener linguistischer Richtungen. Für die Konversationsanalyse gilt die Maxime, dass die Forschung strikt induktiv und korpusbezogen vorgehen muss. Das heisst, dass für die Analyse keine vorgefassten Kategorien (mit festen Funktionszuschreibungen) verwendet werden dürfen; vielmehr soll untersucht werden, wie die Strukturen und Kategorien des Gesprächs durch die Gesprächsteilnehmer gemeinsam erarbeitet werden.⁵ Das heisst, dass die Äusserungen einzelner Sprecher nicht isoliert analysiert werden dürfen (Hakulinen & Selting, 2005a). Auf Grund der Äusserungen, die von den Rezipienten ganzheitlich interpretiert werden, muss der Analytiker rekonstruieren, wie einzelne sprachliche Bausteine systematisch dazu beitragen, das Gemeinte verständlich zu machen (S. 3).

Deppermann (2005) zeigt in seinem Beitrag, wie lexikalische Items abhängig vom Konversationskontext eine spezielle Bedeutung annehmen können. Er zeigt dies am Beispiel der *Darstellung von Kontrasten* in der Konversation, die nicht immer auf der lexikosemantischen Unterschiedlichkeit der verwendeten Ausdrücke beruht, sondern gewissermassen durch beliebige Ausdrücke, die im lokalen Gesprächskontext kontrastive Bedeutung erlangen, zum Ausdruck gebracht werden kann. Zwei Strategien für die Darstellung von Kontrasten, ihre sequentielle Konstruktion und die Auswirkungen für die Rezeption werden an Gesprächsdaten aus verschiedenen Korpora vorgeführt: Einerseits eine *frame-based interpretation*, die einen Gegensatz im gleichen Handlungsrahmen konstruiert, andererseits eine Strategie, die der Autor *maximization of contrast* nennt: Der Hörer realisiert, dass der Sprecher mit zwei verwendeten Ausdrücken einen Kontrast darstellen will und interpretiert diese (unabhängig von deren Kernbedeutung) so, dass der Gegensatz maximiert wird. Bei häufigerer Verwendung kann sich die Kontrastierungsfunktion dieser Gegensatzpaare verselbständigen und längerfristig in einer grammatikalisierten linguistischen Struktur kondensieren (Grammatikalisierung, siehe Kap. 6.8).

⁵Die Autoren räumen allerdings ein, dass syntaktische Strukturen und lexikalische Items und Kategorien kaum als nicht vorgegeben vorgestellt werden können. «Sie sind aber im Interpretationsprozess zu hinterfragen...» (Hakulinen & Selting, 2005a, S. 1).

3.2.2 Gesprochene-Sprache-Forschung

Die Gesprochene-Sprache-Forschung entstand bereits in den 1960er Jahren und wurde 1964 offiziell als Forschungsgebiet in das neu gegründete *Institut für Deutsche Sprache* (IDS) aufgenommen. Sprachbegriff und Sprachstrukturen waren stark geprägt durch schriftliche Texte; kommunikative und dialogische Phänomene galten als Störelemente und wurden deshalb aus den Gesprächsdaten ausgeblendet; die Analyse der gesprochenen Sprache beschränkte sich anfänglich auf die formale und statistische Untersuchung von Syntaxphänomenen (Schwitalla, 2001, S. 897). Im Laufe der 1970er und 1980 Jahre veränderte sich die Gesprochene-Sprache-Forschung unter dem Einfluss der Konversationsanalyse zur linguistischen Gesprächsanalyse (S. 896). Deppermann spricht von einer Renaissance der GSF: «Mit der Interaktionalen Linguistik erleben Fragestellungen der in den 1960er und 1970er Jahren praktizierten *Gesprochene-Sprache-Forschung* unter veränderten Vorzeichen eine Renaissance.» (Deppermann, 2007, S. 19).

Die gesprochene Sprache (siehe Beispielerzählung auf S. 124) weist regelmässig Phänomene und Konstrukte auf, die gemäss Prämissen einer traditionellen Systemlinguistik irregulär sind und die aus der Sicht der traditionellen Grammatiktheorien als unkorrekte Randphänomene und fehlerhafter Sprachgebrauch taxiert werden.

Eine in diesem Sinne praxisorientierte Perspektive auf sprachliche Phänomene impliziert die Abkehr von einem streng systemlinguistischen Verständnis sprachlicher Phänomene in der Tradition des Strukturalismus: Statt Sprache als ein von der Verwendung möglichst unabhängiges System – losgelöst von kommunikativen, sozialen, kulturellen, medialen und funktionalen Dimensionen – zu konzeptualisieren, werden die strukturellen Eigenschaften von Sprache im Gegenteil als Ergebnis ihrer interaktiven Funktionen und Gebrauchsweisen betrachtet. (Günthner & Bückler, 2009, S. 3)

Schwitalla (2003) und Schwitalla und Tiittula (2009) beschreiben spezielle Merkmale der *gesprochenen Sprache*, wie Verarbeitungs- und Produktionsbedingungen und linguistische Merkmale. Für das Sprechen und Hörverstehen gelten besondere Bedingungen, weil das Arbeitsgedächtnis für das Hörverstehen nur ca. 2–3 Sekunden speichern kann; nach dieser Zeit ist der exakte Wortlaut nicht mehr präsent (Schwitalla & Tiittula, 2009, S. 40). Typische linguistische Merkmale, die im schriftlichen Sprachgebrauch nicht üblich sind, sind sukzessive Präzisierungen, Nachträge, Kor-

rekturen, Links- und Rechtsherausstellungen. Typisch für mündliche Verständigung sind auch verblose Sätze, Brüche im Satzbau oder Abbruch eines begonnenen Satzes (Anakoluth), Weglassen eines Buchstabens oder einer Silbe am Ende eines Wortes (Apokope), die Schwa-Elision (Weglassen einer unbetonten Silbe) oder Enklisen (Integration eines Wortes in ein anderes; Bsp. «geht's?»). Weitere Besonderheiten sind Einwortsätze, Wortverkürzungen, spezielle Präfixe (rum-) (Schwitalla, 2003; Schwitalla & Tiittula, 2009). Ausserdem gibt es auch bezüglich Prosodie und bezüglich Lexik und Wortsemantik wichtige Besonderheiten, die die gesprochene Sprache auszeichnen, wie langsames, schnelles, lautes oder leises Sprechen, Akzente, Rhythmus, Tonhöhenregister und -variation, Intonationsverlauf und Pausen. Auf der inhaltlichen Ebene gibt es mehr Orts-, Zeit- und Personendaxis, bildhafte Ausdrucksweise und derbe Wörter, Übertreibungen, vage Bezeichnungen und Vagheitsindikatoren. Für den laufenden Dialog sind Gesprächspartikeln, Gradpartikeln und Interjektionen wichtig, die Funktionen wie Segmentierung, Rückversicherungs- und Rezeptionssignale, Korrekturen, Redeeinleitung und Pausenfüller übernehmen (Schwitalla & Tiittula, 2009, S. 40/41).

Die Existenz von *Korpora gesprochener Sprache* eröffnet auf verschiedenen Forschungsfeldern neue Möglichkeiten, von der Linguistik über die Sozialwissenschaften bis zur Informationstechnologie. Damit stellt sich mehr denn je die Frage, ob wir eine *Theorie der gesprochenen Sprache* brauchen (Günthner, 2007a), oder auch ein *Grammatik der gesprochenen Sprache*. Grammatik wird hier verstanden als eine Art «Institution», d. h. als eine Sammlung von Regeln, die sich als Resultat vergangener kommunikativer Handlungen herausgebildet hat. Auch Auer (2007) plädiert für eine spezielle Syntax für die gesprochene Sprache. Er stellt dabei die *Prozesshaftigkeit* der Mündlichkeit ins Zentrum; mündliche Sprache wird linear in der Zeit produziert und rezipiert, ist dialogisch (die Gesprächsteilnehmer konstruieren die syntaktischen Einheiten gemeinsam). Mündliches Kommunizieren erfordert schnelles Reaktionsvermögen auf die Anforderungen der Gesprächsentwicklung; viele Satzstrukturen sind deshalb musterhaft festgelegt und können als ganze Einheiten abgerufen werden (Vgl. die Untersuchungen zur formelhaften Sprache bei Wray (2000, 2005) im Kapitel 4.7). Der emergente Charakter mündlicher Syntaxkonstruktionen erfordert eine *Online-Grammatik* (Auer, 2007, S. 96), oder eine inkrementelle Syntax. In einer radikal auf Interaktion und Emergenz ausgerichteten Anschauungsweise wird Grammatik schliesslich auf Performanz reduziert, im Gegensatz zur Reduktion auf die Kompetenz bei den Generativisten. Beide Extrempositionen sind wenig nützlich. Für die weitere Arbeit soll der Standpunkt eingenommen

werden, dass es grammatikalisierte und lexikalisierte Strukturen gibt, die von Sprechern schnell und als Ganzheiten abgerufen werden können, dass aber ein kreativer Umgang mit diesen Strukturen in der Kommunikation möglich ist.

3.2.3 Beispiele interaktiver Gesprächspraktiken

Vorgeformtes, Formelhaftes

Gülich (2007b) verwendet den Begriff «Vorgeformtheit», um auf eine spezielle und individuelle Verwendung von Redewendungen und Idioms in Gesprächen hinzuweisen. Erzähler schaffen sich damit einen eigenen Sprachstil; Gemeinplätze, Idiome und Metaphern werden in einer für den Sprecher eigenen Weise verwendet. Der Rekurs auf Vorgeformtes zeigt eine Orientierung an Modellen. Ansatzweise ist die Verwendung von vorgeformten Redeteilen störungsspezifisch. Vorgeformte Strukturen übernehmen spezielle Rollen im Formulierungsprozess und im Gespräch.

Gemäss Gülich (2007b) wird Vorgeformtes nicht einfach reproduziert, sondern die vorgeformten Elemente werden bearbeitet und an den Kontext angepasst. Sie wertet den Rekurs auf Vorgeformtes als eigenständiges und anspruchsvolles *Formulierungsverfahren* und nicht als Entlastung von der Formulierungstätigkeit. Im Gegensatz dazu sehen Günthner und Bückler (2009) das Formelhafte vor allem als Vereinfachung der Kommunikation unter Zeit- und Handlungsdruck. Unter den Produktions- und Rezeptionsbedingungen der gesprochenen Sprache sind rekurrente und verfestigte Sprachmuster unabdingbar. Kapitel (4.7) stellt formelhafte Sprache (formulaic language) ausführlicher vor.

Rhetorische Verfahren und sprachliche Marker

Günthner (2006a) beschreibt die Kommunikation von *Panikattacken*. Rhetorische, grammatische und prosodische Formen vermitteln den Ausdruck der spezifischen Panikerlebnisse und körperlichen Extremerfahrungen. Dazu verwenden die Erzähler Kontrastierungsverfahren und narrative Verfahren wie (‹dichte Konstruktionen›, Verbspitzenstellung, narratives Präsens, subjektlose Infinitkonstruktionen, ‹minimale Setzungen›), oder sie charakterisieren die Episoden als unbeschreibbar. *Infinitkonstruktionen* (Bsp. «Ich sofort angehalten...») reduzieren die Satzinformationen auf zwei Angaben: auf den Ereignisträger und das Ereignis. Eine noch reduzierte Form findet sich in den *subjektlosen Infinitkonstruktionen* (als Fortsetzung des vorherigen

Beispiels: «Ausgestiegen...»). Infinitkonstruktionen sind Ressourcen szenischer Performanz und dienen im interaktiven Gebrauch der Spannungserzeugung und dramatischen Darstellung. D. h. Konstruktion *und* kommunikativer Kontext sind wichtig (Günthner, 2006b, S. 9). Vorwurfsäußerungen als kommunikative Aktivität bewirken in der sequentiellen Organisation Folgeaktionen wie Erklärungen oder Gegenvorwürfe (Bsp. Bauplan von <warum>-Vorwürfen: Zuschreibung von Verantwortung, Negation von Vernünftigkeit) (Günthner, 2006b, S. 13). Pragmatische Aspekte der Situation und des Kommunikationskontextes bekommen eine wichtige Bedeutung (Gesprächstyp). Für das vorliegende Projekt eines Konstruktionslexikons stellt sich wieder die praktische Frage und Aufgabe: Was davon kann man lexikalisieren?

Einige Beispiele von Konstruktionen als sprachliche Marker sollen das bisher Gesagte illustrieren und gleichzeitig einen Blick auf die später beschriebenen Überlegungen für eine praktische Umsetzung im Konstruktionslexikon ermöglichen:

Irreguläre Ausdrücke: Das Beispiel «typisch Frau» kann durch das Muster (die Konstruktion) <Unflektiertes Adjektiv + Nomen> repräsentiert werden. Das Adjektiv ist festgelegt, das Nomen variabel. Dieser Ausdruck wird als eigenständiger Eintrag ins Lexikon übernommen.

Passivkonstruktionen: Deppermann (2006a, S. 51): Aktiv- und Passivkonstruktion sind eigenständig und nicht durch syntaktische Transformationen ableitbar, weil sie unterschiedliche semantische, pragmatische und diskursive Eigenschaften besitzen (z. B. unterschiedliche semantische Rollen).

Diskursmarker: Imo (2007) beschreibt Diskursmarker in mündlichen Narrativen, Norrick (2001) beschreibt Funktionen der englischen Diskursmarker «well» und «but».

Füllwörter und -phrasen: Im Rahmen einer korpusbasierten Analyse von phraseologischen Ausdrücken untersuchen Maclagan, Davis und Lunsford (2008) den speziellen Sprachgebrauch von *Alzheimer Patienten*. Die Autorinnen zeigen, wie die Verwendung von Phraseologismen verschiedener Art (vorgefertigte Muster, fixe Formeln, Füllwörter- und phrasen (extenders) und metonymische Ausdrücke) den Alzheimer-Patienten ermöglicht, den Anschein von Sprachkompetenz aufrecht zu erhalten. Im Verlauf der Krankheit werden die Formeln weiterhin verwendet, sind aber zunehmend nur noch für die Konversationspartner

der Kranken verständlich, weil neben den Formeln die Sprachkompetenz ständig abnimmt. Trotzdem ermöglichen sie den Patienten, sich als kompetente Interaktions- und Konversationsteilnehmer zu positionieren.

Konstruktionen werden ausführlicher im Kapitel (6.2) beschrieben.

3.3 Bedeutungskonstitution

Im Folgenden beziehe ich mich auf den Begriff der *Bedeutungskonstitution* wie ihn Deppermann (2006e) beschreibt, und auf den Begriff der *meaning construction* (Radden, Köpcke, Berg & Siemund, 2007). Der Begriff der *Bedeutungskonstitution* und ein Theoriegebäude dazu existieren bis anhin nicht in einer allgemein anerkannten Weise. Im hier intendierten Sinn bezieht sich der Begriff immer auf konkrete Handlungskontexte, also auf den Hintergrund tatsächlicher Sprachverwendung. Die *Bedeutung* von Wörtern, Ausdrücken und Sätzen ist nicht kontextfrei vorgegeben, sondern sie wird in einem bestimmten Kontext und in einem zumeist interaktiven Prozess erst hergestellt.⁶

Bedeutungskonstitution umfasst alle sprachlichen Ebenen wie Lexikon, Syntax, Wort- und Satzsemantik, Pragmatik, sowie auch die Text- und Diskursebene (Deppermann, 2007, S. 211). Sinn und Bedeutung entstehen in der sprachlichen Kommunikation nicht auf der Wortebene, sondern erst in grösseren Einheiten wie Ausdrücken oder Sätzen. Dies steht im Gegensatz zum Fregeschen Prinzip der lexikosyntaktischen Kompositionalität der Satzbedeutung. Das *Kompositionalitätsprinzip* wirft verschiedene ungelöste Fragen und Probleme auf (nach Deppermann, 2006e):

- Wie funktioniert die lexikosemantische *Disambiguierung* zwischen mehreren potentiellen Wortbedeutungen, wenn Wortbedeutungen ohne Kontext betrachtet werden?

⁶«Kontrovers wurde vor allem der zu Grunde zu legende Bedeutungsbegriff diskutiert, der von Lexikologen und Kognitionslinguisten zumeist erheblich enger gefasst wird als von Gesprächsforschern. Umstritten blieben auch der Nutzen und die methodische Verifizierbarkeit des Rekurses auf mentale Kategorien bei der empirischen Untersuchung. Weitgehende Einigkeit ergab sich dagegen in der Frage der Notwendigkeit, Fragen der Kontextabhängigkeit systematischer und zentraler als bisher üblich in linguistischen Bedeutungstheorien zu verankern und dabei den dialogischen Prozessen der Bedeutungsspezifikation größeres theoretisches Gewicht zu verleihen und sie verstärkt zum empirischen Untersuchungsgegenstand zu machen.» (Bericht über die AG 9 «Bedeutungskonstitution», A. Deppermann, 2000), <http://www.gespraechsforschung-ozs.de/heft2000/tb-dgfs2000-ag9.pdf>

- Wie funktioniert die referenzielle Interpretation, d. h. wie soll der *Referent* bestimmt werden, auf den sich ein Ausdruck bezieht?
- Viele Sätze sind unvollständig oder *unterspezifiziert* und haben doch erkennbare Bedeutung. Wie ist dies zu erklären?
- Die Konzeption einer *wörtlichen Bedeutung* ist empirisch unhaltbar (vgl. Gibbs, 1994).
- Und als Quintessenz folgert Deppermann (2006e, S. 16): «Die lexikalische Bedeutung hat gar keine kontextfreie Existenz».

Gegen das Prinzip der lexikosemantischen Kompositionalität stehen verschiedene Annahmen, unter anderem aus Gesprächsforschung und interaktionaler Linguistik, wie sie bereits in vorherigen Kapiteln angesprochen wurden (Deppermann, 2006e):

- Wortbedeutungen sind *Bedeutungspotenziale*, die nur kontextgebunden als kommunikativ relevant spezifiziert werden.
- Bedeutungskonstitution ist ein *soziales, öffentliches Geschehen* (und nicht nur eine kognitive Leistung⁷).
- Bedeutungskonstitution erfolgt über Prozesse der *interaktiven Aushandlung* und des *recipient design*, d. h. Formulierungen werden dem Hörer und dessen individuellen Handlungsinteressen und Vorwissen angepasst.
- Im Gespräch jeweils aktuelle und *saliente Themen* beeinflussen die Herstellung von Bedeutung.

Bedeutung wird laufend vom Sprecher generiert und vom Hörer entsprechend verstanden, indem dieser die aktuelle Bedeutung aus den potentiell möglichen und ihm bekannten erschliesst. Ein praktisches Beispiel dafür, wie dies ablaufen kann, liefert Franceschini (2006), die Verständigungspraktiken von nicht italienisch sprechenden Deutschschweizern mit Italienern beschreibt. Die Sprecherinnen erschliessen die Bedeutung der Äusserungen aus dem Kontext und aus der Gesprächssequenz.

Linell (2009) bespricht ausführlich die *Bedeutungskonstitution* und weist darauf, dass Bedeutung und Herstellung von Bedeutung zentrale psychologische Konzepte darstellen (Bruner, 1990). Bedeutung entsteht in der

⁷Selbstverständlich spielen kognitive Prozesse bei der Herstellung von Bedeutung ebenfalls eine grosse Rolle; Deppermann (2006e, S. 26) plädiert für eine integrierte Betrachtung von kognitiven und interaktiven Prozessen.

situieren, d. h. von der Situation abhängigen Interaktion und im Diskurs. Linguistische Ressourcen haben nicht Bedeutung per se, sondern eher *Bedeutungspotenziale*. Ausserdem sind die Bedeutungen, die Gesprächsteilnehmer konstruieren, theoretisch nie abgeschlossen; sie können immer weiter entwickelt und spezifiziert werden. Allerdings räumt Linell ein, dass aus praktischen Gründen die Herstellung von Sinn (sense-making) auch einmal abgeschlossen werden muss (S. 224).⁸

Ein weiteres Merkmal der Bedeutungskonstitution ist die *Implicitness*.⁹ Ein grosser Teil der Herstellung von Bedeutung in der Interaktion geschieht *implicit*; es könnten gar nicht alle zu diesem Zweck ablaufenden Prozesse explizit ausgedrückt werden. Auch die Kommunikation basiert zu einem grossen Teil auf impliziten Annahmen (Linell, 2009, S. 224). Den gleichen Punkt betonen Radden et al. (2007) und brauchen dafür den Begriff der *Unterspezifizierung*, wovon die Implicitness ein Ausdruck ist. Unterspezifizierung ist ein *kommunikatives Prinzip*¹⁰ und bei der Herstellung von Bedeutung ein wichtiger Faktor. Jede Äusserung ist in dem Sinne unterspezifiziert, dass sie nicht alle Aspekte ausdrückt, die für eine vollständige Interpretation nötig wären. Entsprechend lautet die Definition von *meaning construction* bei Radden et al. (2007):

Meaning construction is an on-line mental activity whereby speech participants create meanings in every communicative act on the basis of underspecified linguistic units. (S. 3)

Radden et al. (2007) postulieren drei Typen von Unterspezifizierung, die im Folgenden kurz beschrieben werden.

3.3.1 Unterspezifizierung durch Implicitness

Die Unterspezifizierung beruht hier darauf, dass es für die intendierte Bedeutung keine entsprechende (spezifische) linguistische Einheit gibt, oder lediglich eine ohne spezifische Bedeutung (Radden et al., 2007, S. 5). Einzelworte und Ausdrücke sind assoziiert mit einer oder meistens mehreren kon-

⁸ «sometimes we must rest and go to sleep, eventually we die.» (Linell, 2009, S. 224).

⁹ durch «Mitinbegriffensein» oder «Selbstverständlichkeit» nur unzureichend übersetzt.

¹⁰ Hintergrund der Unterspezifizierung ist der Begriff der «Bedeutungslosigkeit» (meaninglessness). Radden et al. (2007, S. 2) nennen als Beispiel eine rechteckige Figur die Ähnlichkeit mit einem Rahmen hat; die vier Ecken der Figur verbinden jedoch die Rahmenkanten nicht auf sinnvolle Weise; die Figur ist bedeutungslos. In der experimentellen Aufgabe, diese Figur aus dem Gedächtnis nachzuzeichnen, machen jedoch die Versuchspersonen in der Regel eine sinnvolle Figur daraus, z. B. einen Rahmen mit korrekt verbundenen Ecken.

ventionalisierten Bedeutungen, die in einem bestimmten Kontext weitere, speziellere Konzeptualisierungen aktivieren. Ein Beispiel (3) aus (Radden et al., 2007) mit dem Frame *«commercial event»* soll dies illustrieren: Die Verben *«sell»* und *«buy»* aktivieren je einen bestimmten Kontext, die weiteren Framekomponenten setzen spezifische Prozesse der Präzisierung in Gang (oder führen auf einen andern Frame). Dies bewirkt, dass die Höhe des Geldbetrages in den beiden Beispielen unterschiedlich eingeschätzt wird.

- (3) a. Harry sold his car for a good price.
- b. Harry bought his car for a good price.

Als zweites Beispiel erwähnen Radden et al. (2007, S. 6, nach Fauconnier & Turner): Der Ausdruck *«safe»* (sicher), z. B. im Satz *«etwas ist sicher»* löst ein Gefahrenszenario aus, das in der lexikalischen Einheit nicht a priori präsent ist. Es erfolgt eine Aktivierung von verschiedenen Konzepten oder *mental spaces* (conceptual blending).

Bei komplexen Ausdrücken resultiert die Unterspezifizierung daraus, dass die einzelnen Komponenten des Ausdrucks ihre Bedeutung im Ausdruck teilweise verlieren. Unterspezifizierung besteht sowohl für die einzelnen Komponenten des Ausdrucks als auch für die Kompositionsbedeutung (Beispiel *«Rollstuhl»*: die einzelnen Teile des Kompositums sagen nichts über den Verwendungszweck des Gegenstandes aus).

3.3.2 Unterspezifizierung durch Nicht-Determiniertheit

Die Nicht-Determiniertheit (indeterminacy) könnte auch mit *Vagheit* bezeichnet werden. Auch dazu soll ein Beispiel (4) aus Radden et al. (2007, S. 6) gezeigt werden:

- (4) a. The heat broke the guitar.
- b. A falling radio broke the guitar.

Unser Weltwissen sagt uns, dass die beiden *«break»*-Vorgänge zwar etwas Ähnliches bezeichnen, aber dass die Qualität der Aktivität *«break»* in den beiden Beispielen ganz unterschiedlich ist, z. B. in Bezug auf die Dauer und die physikalische Intensität des Geschehens.

Linell (2009, S. 226) betont, dass *Vagheit* (vagueness) nicht einen Mangel in der situierten Bedeutungsherstellung bedeutet, sondern dass Vagheit eher die Voraussetzung bildet, um projektive und rekonstruktive Aufgaben in der Interaktion (in kommunikativen Projekten) zu lösen (diese Ansicht

widerspricht allerdings den kognitivistischen Annahmen zielgerichteten Verhaltens).

3.3.3 Unterspezifizierung durch Inkompatibilität

Damit bezeichnen die Autoren *meaning construction* als Konfliktlösung zwischen *inkompatiblen Konzepten*. Dieser Typ kommt nur in zusammengesetzten Ausdrücken vor und ist die häufigste Art der Unterspezifizierung. Inkompatibilitäten zwischen Kernbedeutungen von Einzelkomponenten eines Ausdrucks gibt es sowohl auf lexikalischer als auch auf grammatischer Ebene. So können z. B. Nomen mit widersprüchlichen Adjektiven versehen werden («the slow fast dog»), zu dieser Strategie zählen auch *Oxymoron* und *contradictio in adjecto* («der stumme Schrei», «Minuswachstum», «der alte Knabe», «screaming silence»¹¹).

Lexiko-grammatische *Inkompatibilität* liegt im bekannten Beispiel von Goldberg (1995) vor: «Fred sneezed the napkin off the table». Unterspezifiziert ist in diesem Beispiel nicht die Konstruktion, sondern das Lexem «sneeze» (niesen), welches in dieser Konstruktion (CAUSED-MOTION) eine unübliche Valenz erhält (ditransitiv statt intransitiv).

Meaning construction betrifft nicht nur «neue» Bedeutung; d. h. nicht nur neuartige Bedeutung wird konstruiert, sondern jede Bedeutung in jeder Situation wird bis zu einem gewissen Grad neu ausgehandelt. Die Prozesse der Bedeutungsherstellung betreffen sowohl Produzenten als auch Rezipienten von sprachlichen Äusserungen. Das heisst, dass im Kommunikationsvorgang zweimal Bedeutung hergestellt wird (beim Meinen und beim Verstehen), zweimal eine iterative Prozedur, die gelingen oder scheitern kann.

3.3.4 Bedeutungskonstitution und Konstruktionsgrammatik

Zur Frage, ob Bedeutung quasi kontextfrei in lexikalisierten Mustern vorhanden ist oder ob im Gegensatz dazu Bedeutung nur interaktiv entsteht und gar nicht konserviert werden kann, machen sich Fried und Östman (2005) im Rahmen der Konstruktionsgrammatik die folgenden Überlegungen (vgl. Kap. 6.5):

¹¹ *Oxymoron*: rhetorische Figur, bei der eine Formulierung aus zwei gegensätzlichen, einander widersprechenden oder sich gegenseitig ausschliessenden Begriffen gebildet wird (Quelle: Wikipedia).

Construction Grammar does not assume that meaning is only negotiated in actual interaction. Instead, lexical and phrasal meanings can be specified, in a given construction, as part of the interactants knowledge of language. (Fried & Östman, 2005, S. 1755).

Bedeutung wird nicht nur in praktischer Interaktion verhandelt, sondern die lexikalische und phrasale Bedeutung (von Konstruktionen) kann auch als Ausdruck des Sprachwissens der Sprecher/Interaktanten verstanden werden.

Gemäss Fried und Östman (2005, S. 1756) unterscheiden sich Konversationsanalyse und Konstruktionsgrammatik in ihren Ansichten bezüglich *Variabilität* der Sprache und *Sprachwandel*. CA beschäftigt sich mit dem Entstehen von Diskursmustern im Gespräch, im Dialog und als gemeinsame Leistung der Gesprächsteilnehmer, während sich die Konstruktionsgrammatik eher mit festgelegten Mustern und statischen Konstruktionen beschäftigt (Imo, 2007), die zwar eine gewisse Variabilität haben, welche aber explizit als Variation des Musters beschrieben werden kann (Fried & Östman, 2005, S. 1756). Konversationsanalytische Studien haben gezeigt, dass Diskursmuster von den Teilnehmern gemeinsam produziert werden können. Deshalb kann man über die *Ko-Konstruktion* von Bedeutung und Verstehen im Diskurs sprechen und gleichzeitig akzeptieren, dass das Resultat des Ko-Konstruierens mindestens teilweise einer Kombination der formelhaften Redewendungen entspricht, die im mentalen Lexikon der Sprecher gespeichert sind (Fried & Östman, 2005, S. 1775).

3.3.5 Emergent Grammar

Als Begründer der Emergent Grammar gilt Paul Hopper (Hopper, 1987). Grammatik wie Sprache sind soziale real-time Phänomene und deshalb prozesshaft, ihre Strukturen sind immer unterwegs und kommen nie an. Emergenz meint, dass Strukturen und Regeln aus dem Diskurs entstehen und auch den weiteren Diskurs wieder formen, d. h. die Strukturen sind Epiphänomene des Sprachgebrauchs. Hopper charakterisiert die Emergent Grammar mit folgenden Worten:

The notion of Emergent Grammar is meant to suggest that structure, or regularity, comes out of discourse and is shaped by discourse as much as it shapes discourse in an on-going process. (Hopper, 1987, S. 2)

Vorausgegangene *Äusserungen* bilden die Basis für neue (sowohl die nächsten im Ablauf als auch neuartige) *Äusserungen*. Dieser Aspekt der Sprache wird von Hopper vorgeformt oder vorgefertigt genannt (pre-patterned, pre-fabricated): im spontanen Diskurs komponieren und sprechen wir gleichzeitig (Hopper, 1987, S. 4). Hopper etabliert das Konzept der *Figur*, eine Phrase oder ein kurzer Satz, der hoch standardisiert ist, eine rudimentäre interne Struktur hat und näher bei einer Formel ist als bei einem frei generierten Satz. *Diskurs* besteht zu einem grossen Teil aus einer Aneinanderreihung solcher Figuren (Hopper, 1987, S. 6). Je nützlicher sich eine Konstruktion im Sprachgebrauch erweist, desto mehr wird sie im Gebrauch strukturiert und verfestigt werden, indem sie Konsistenz unabhängig vom Kontext bekommt und so als Basis für weitere Variationen und Erweiterungen dient (Hopper, 1987, S. 7). Ein ähnliches Konzept, wenn auch von einer andern Seite her, verfolgt Hoey (2005, S. 7): Die Bedeutung eines Wortes entwickelt sich durch seine Rezeptionsgeschichte (geschrieben und gesprochen). Das Wort wird durch Kontext und Kotext semantisch aufgeladen. Dasselbe gilt über das Wort hinaus auch für Ausdrücke und Sätze.

3.4 Satzsemantik – Pragmatik

Die oben angesprochenen Fragen der Bedeutungskonstitution und der Aushandlung von Bedeutung bringen es für die erzähl- und gesprächsanalytische Untersuchung von Gesprächen mit sich, dass nicht nur die linguistische und wortsemantische Ebene für die Interpretation betrachtet werden darf, sondern dass die im Gebrauch relevanten Sprachmerkmale auf die *Satz- und auf die Diskurssemantik* sowie auf die *pragmatischen* (und damit ausser-sprachlichen) Merkmale erweitert werden müssen.

Zum Verhältnis der Begriffe *Semantik* und *Pragmatik* beschreibt Polenz (2008) die Hintergründe und die Entstehung dieser Begriffe. Die meisten älteren linguistischen Theorien hatten Sprache auf ein autonom funktionierendes System reduziert. Durch die Konzeption der Sprache als *Handeln* bekommen die Vorgänge des *«Meinens»* und *«Verstehens»* eine wichtige Bedeutung (Hörmann, 1988). Karl Bühler unterscheidet in seiner Sprachtheorie die *Darstellungsfunktion* (Sprache als Symbol) und die *Ausdrucksfunktion* (Sprache als Symptom für Absichten, Gefühle, Stimmungen), und – auf den Empfänger/Hörer bezogen – die *Appellfunktion* der Sprache (Sprache als Signal). Durch die Einführung von Ausdrucks- und Appellfunktion kann Bühler als Begründer der *Pragmatik* bezeichnet werden. Auch Polenz (2008, S. 68) postuliert, dass Semantik durch Pragmatik erweitert und neu

konzipiert werden muss.

3.5 Situation – Gesprächssituation

Der Begriff der *Situation* ist nicht eindeutig, es gibt dazu viele und heterogene Konzepte. Verbreiteter ist der Ausdruck *Kontext*. Das gleiche oder damit verwandte Phänomene bezeichnen aber auch die Begriffe Handlungsschema, Rahmen (frame), Gesprächstyp, Interaktionstyp oder Kommunikationstyp. Situations- und Kontextbegriff wurden von Bronislaw Malinowski und John Rupert Firth als Komponenten der Gesprächssituation eingeführt und mit der funktionalen Sicht verknüpft, wonach spezifische sprachliche Formen situationsspezifische Funktionen haben. «Situationen stehen also im Spannungsfeld von wiederholbarer Kategorisierung und biographischer und historischer Einzigartigkeit.» (Deppermann & Spranz-Fogasy, 2001, S. 1150).

Für das vorliegende Lexikonprojekt sind die Merkmale *Situation* und *Kontext* wichtig, bieten aber die grössten Probleme bei der Umsetzung in ein praktisch anwendbares Lexikon. Situation und Kontext können nur bedingt als Parameter in das Lexikon einfließen, da sie zum grössten Teil im Rahmen des Diskurses verankert sind. Die pragmatischen Merkmale im neuen JAKOB-Lexikon (Domain, Texttyp, Register und Topic, siehe Kap. 9.1.3) sind der Versuch, die Eigenschaften von Situation und Kontext in der lexikographischen Darstellung zu integrieren.

Die *Gesprächslinguistik* greift zunehmend auf Konzepte der Pragmatik zurück, mit einer «Wendung vom Satz zur Äusserung, von der grammatischen Kompetenz zum sprachlichen Handeln, vom linguistischen System zum Kommunikationsereignis» (Deppermann & Spranz-Fogasy, 2001, S. 1148). Dies steht im Gegensatz zur in der Systemlinguistik vertretenen Sicht des Kontextbegriffs als einer unerklärten Restkategorie (S. 1149). Deppermann und Spranz-Fogasy (2001, S. 1151f) zeigen, wie das Konzept der *Situation* in verschiedenen Disziplinen aufgenommen wird:

- Im so genannten *Variablenmodell der Gesprächssituation* ist die Situation eine unter anderen unabhängigen Variablen, die im Gespräch identifiziert und gemessen werden können (z. B. syntaktische Merkmale, lexikalische Wahlen etc.). Typischer Vertreter dieser Richtung ist z. B. Labov (1972).
- Sprechakttheoretische Ansätze definieren Sprechakte als sprachliche *Handlungen*, mit denen bestimmte Ziele erreicht werden sollen. Sprechakttypen implizieren dabei bestimmte Situationskonstellationen.

- Im Britischen *Kontextualismus* wird dem <context of situation> grosse Beachtung geschenkt; der Kontext bestimmt die Varietäten des Sprachgebrauchs (Register). In der *systemisch-funktionalen Linguistik* (Halliday & Matthiessen, 2004) werden Aussagen zum Situationsbegriff als wichtige linguistische Parameter eingeführt (z. B. drei situative Dimensionen: *field, tenor, mode*). Situation und Register werden später von Biber und Finegan (1994b) aufgegriffen (*dialect, register, genre*).
- Kognitionswissenschaftliche Ansätze greifen die Konzepte von Situation und Kontext ebenfalls auf, so beispielsweise Dijk (2008) in seinem *Kontextmodell*. Kontextmodelle sind die persönlichen Interpretationen (Relevanzsetzungen) des sozialen Umfeldes durch die Interaktionsteilnehmer und bestimmen weitgehend die individuelle Diskursproduktion und -rezeption.
- Dell Hymes modelliert situative Faktoren im so genannten SPEAKING-Modell (siehe Kapitel 7.3.2).
- Gemäss *Interaktionaler Soziolinguistik* ist die Gesprächssituation nicht nur vorgegeben, sondern wird auch im Gespräch selbst hergestellt (durch Kontextualisierungshinweise wie nicht-propositionale, verbale, para- oder nonverbale Phänomene, Formeln, Kode-Wahl, Prosodie, Mimik, Gestik etc.). D. h. diese Merkmale sind nicht wie *Register* statische, an Situationen gebundene Marker, sondern flexible Inferenzhinweise und kommen deshalb fast nur in Einzelfallbeschreibungen vor.
- Im Verständnis von *Ethnomethodologie* und *Konversationsanalyse* ist jede Interpretation situationsgebunden, Situation und Kontext werden durch Äusserungen indiziert (kommunikationstypspezifische Handlungsschemata). Die Gesprächssituation wird durch Gesprächsaktivitäten hergestellt und durch Aufzeigeleistungen verdeutlicht. Jeder Gesprächsbeitrag ist zugleich kontextgebunden und kontexterneuernd.

In der Gesprächslinguistik wird *Situation* oft als eine Art Restkategorie eingesetzt, um Äusserungs- und Diskurseigenschaften zu charakterisieren, die nicht mit den zur Verfügung stehenden methodischen Mitteln beschrieben werden können oder die nicht dazu passen (Deppermann & Spranz-Fogasy, 2001, S. 1157).

Die Frage, ob Situation und Kontext *statisch* oder *dynamisch* zu verstehen sind, ist wichtig und illustriert die Schwierigkeiten bei der Formulierung

von Lexikoneinträgen und bei der Umsetzung der Kodierprozeduren in der Erzählanalyse JAKOB. Dazu (Deppermann & Spranz-Fogasy, 2001):

Die skizzierten Differenzen verdichten sich in der Frage, ob Gesprächssituationen statisch oder dynamisch sind. Im Variablenmodell und im britischen Kontextualismus werden Situationen als Mikro-Realisierungen stabiler makrosoziologischer Rahmenbedingungen und psychologischer Dispositionen verstanden; sie sind folglich durch Gesprächsaktivitäten nur sehr begrenzt zu verändern. Die interaktionale Soziolinguistik und die Ethnomethodologie gehen dagegen von einem dynamischen Situationsmodell aus, nach dem die Situation im Gespräch stets aufs neue zu bestimmen und zu verändern ist. (S. 1158)

Als Fazit fordern Deppermann und Spranz-Fogasy (2001, S. 1159) die Erarbeitung von Kriterien für die Darstellung und Validierung von Situations- und Kontextwissen als wichtige Verbesserungen für alle Bereiche der Gesprächslinguistik.

3.6 Fazit zum Kapitel Gesprächsforschung und Interaktionale Linguistik

Gesprächsforschung, Konversationsanalyse und Interaktionale Linguistik setzen teilweise andere Schwerpunkte bei der Analyse von Gesprächssequenzen als die im Kapitel zur Narrativik beschriebenen Verfahren. Letztlich ist aber eine scharfe Trennung zwischen narrativen und gesprächsanalytischen Zugängen weder möglich noch wünschbar. Die hier vorgenommene Aufteilung ist deshalb etwas willkürlich.

Interaktion und *Kommunikationsumstände*, auch als *soziale Praktiken* bezeichnet, sind in der Gesprächsanalyse die zentralen Themen. Das Interesse richtet sich grundsätzlich auf das ganze Gespräch, nicht nur auf einzelne Kommunikationsformen wie Erzählungen. Der gesprächsanalytische Zugang ist aus dieser Sicht universaler, wird dem Gespräch als Ganzes eher gerecht. In der Erzählanalyse besteht das Problem, dass unter Umständen für gewisse Erzählerinnen typische und entscheidende kommunikative und psychische Vorgänge sich nicht in Erzählungen sensu JAKOB niederschlagen oder nicht in Erzählungen des gewünschten episodischen Typs, so dass hier eine gewisse Beschränkung der Erzählanalyse festzustellen ist. In diesem Sinne werden die breiter angelegten gesprächsanalytischen Zugänge der Vielfalt

der Patienten und Äusserungsmöglichkeiten in Therapiegesprächen eher gerecht, indem beispielsweise auch formale Merkmale der Gesprächssituation, spezielle Diskursmarker oder formelhafte Ausdrücke beachtet werden.

Ergeben sich daraus neue Aspekte für die Entwicklung des Lexikons? Können allenfalls interaktionelle und konversationelle Funktionen von Äusserungen im Lexikon abgebildet werden? Die neu eingeführten pragmatischen Lexikonmerkmale zielen in diese Richtung (siehe Kapitel 9.1.3). Das Lexikon soll grundsätzlich sprachliche Ausdrücke beschreiben und vor allem unterscheiden können. Der Grundgedanke ist also, dass zwei Lexikoneinträge durch die Gesamtheit der Ausprägungen ihrer Merkmale unterschiedlich sind. Die Kodierung des Lexikoneintrags gemäss Erzählanalyse JAKOB (Kapitel 2.3.4) ist dieser Unterscheidung nachgeordnet; das Unterscheidungspotenzial kann, aber muss nicht durch die JAKOB-Codes ausgedrückt werden.

Wichtig für die Erzählanalyse JAKOB sind die Überlegungen zu *Kontext* und *Situation*. Die Analyse einer Erzählung hat oft wenig Bezug zum grösseren Kontext, wenigstens nicht auf eine in der Anwendung formalisierte Art und Weise. Das Kontextwissen fliesst eher implizit durch die Auseinandersetzung des Erzählanalytikers mit dem ganzen Transkript ein und kann deshalb nicht klar veranschlagt werden. Für das *Lexikon* wichtiger ist der Kontextbegriff im engeren Sinn in Bezug auf die Kodierung (wie beeinflusst der engere Kontext die Kodierung eines Ausdrucks?). Dieser Kontextbezug der Kodierung ist bis jetzt zu wenig explizit reflektiert worden und erst durch die Anwendung des Lexikons in den Fokus der Aufmerksamkeit geraten.

Daran schliessen sich die Fragen an, die im Kapitel zur *Bedeutungskonstitution* dargelegt werden. Wo und wie entsteht Bedeutung beim Meinen und Verstehen? Die eine extreme Sichtweise besagt, dass Bedeutung aus gespeicherten Listen (im mentalen Lexikon) abgerufen wird, die gegensätzliche extreme Sichtweise besagt, dass Bedeutung nur ad hoc in der Situation entsteht. Beide extremen Sichtweisen müssen meiner Ansicht nach verworfen werden zugunsten eines Modells, das Bedeutungspotenziale im mentalen Lexikon annimmt, die durch frequenten Gebrauch in der Kommunikation verfestigt werden, deren Bedeutung aber in der Situation verändert und angepasst werden kann. Weitere Aspekte der *dynamischen Bedeutungskonstitution* werden in den folgenden Kapiteln zur Phraseologie und zur Konstruktionsgrammatik aufgegriffen.

Besonders vielversprechend und zukunftsweisend scheint mir die Kombination von interaktionalen und kognitiven Sichtweisen, wie sie zum Beispiel Deppermann (2006a) für die Zusammenarbeit von Interaktionaler Linguis-

tik und funktionalen Sprachtheorien (Konstruktionsgrammatik) befürwortet.

4 Phraseologie

4.1 Einführung und Überblick

Während in den Kapiteln zum Narrativ und zur Gesprächsforschung kommunikative und interaktive Aspekte des Sprachgebrauchs dargestellt wurden, geht das folgende Kapitel auf die praxisbezogenen linguistischen Fragestellungen zur Verwendung von Ausdrücken oder *Phraseologismen* ein, wie sie unter dem Begriff der *Phraseologie* zusammengefasst werden können. Narrativik und Gesprächsanalyse zeigen, dass Gespräche nicht in Einzelworte zerlegt werden können, sondern dass grössere Einheiten (mindestens «Ausdrücke», Feilke, 1996, S. 66) bis hin zum Diskurs betrachtet werden müssen, um sprachliche Regelhaftigkeiten und die Konstruktion von Bedeutung beim Meinen und Verstehen zu beobachten. Die Phraseologie beschäftigt sich auf einer eher formal-linguistischen Ebene mit Ausdrucksbedeutung und mit *Phraseologismen* oder *Mehrwortausdrücken* (englisch multiword expressions, MWE, oder multiword units, MWU), die von einfachen Wortkombinationen bis zu ganzen Sätzen reichen. Phraseologie wird als «the study of the structure, meaning and use of word combinations» bezeichnet (Cowie, 1994, zitiert nach Granger & Paquot, 2008, S. 27) und befasst sich mit den «festen Wortverbindungen» (Burger, 2004). Das Spektrum der Phraseologie ist sehr weit gefächert und reicht von der Beschreibung von sprachlicher Formelhaftigkeit, wie sie zum Beispiel im Sammelband von Wirrer (1998) erörtert wird, bis zur detaillierten Beschreibung und Klassifizierung der beteiligten Phänomene, wie Mehrwortausdrücke, Kollokationen, Idiome, Metaphern und Metonymie, deren Stellenwert und Funktionen innerhalb der Phraseologie in den entsprechenden Kapiteln dargestellt wird.

Anknüpfungspunkte zur Phraseologie finden sich in der *Computerlinguistik*; phraseologische Fragestellungen in der Computerlinguistik haben mit der Existenz von grossen elektronischen Textkorpora stark an Bedeutung gewonnen und werden im Kapitel (4.4) vorgestellt, zusammen mit der computerlinguistischen Teildisziplin *Word Sense Disambiguation* (WSD), die seit längerer Zeit existiert und versucht, das Disambiguierungs-

problem des Wortsinns mit statistischen Clustermethoden zu lösen. Auch aus der Sicht der *Psycholinguistik* gibt es interessante Anknüpfungspunkte zur Phraseologie, die vor allem die kognitive Verarbeitung bei der Produktion und Rezeption von Wortverbindungen betreffen.

Phraseologie wurde lange Zeit auf der Grundlage von tradiertem Fachwissen und Intuition betrieben; zunehmend gewinnt aber eine korpusbasierte Sichtweise an Bedeutung, die sich am realen Sprachgebrauch orientiert (siehe Kap. 5). Dabei beschränkt sich aber der Begriff der Phraseologie nicht nur auf die Gesprochene Sprache, sondern ebenso auf schriftlich fixierte Texte.

Burger (2007) umreißt die Grundbegriffe der Phraseologie und entwirft eine Terminologie. Phraseologismen bestehen aus Komponenten und sind *feste Wortverbindungen*; dies im Gegensatz zu *freien Wortverbindungen*. Burger unterscheidet zwischen wörtlicher und *phraseologischer Bedeutung*, letztere bezeichnet die übertragene oder auch figurative Bedeutung; der Begriff ist aber umfassender und enthält keine Vorannahmen über die spezifische Art der jeweiligen phraseologischen Eigenart (Burger, 2007, S. 13). Die *freie Bedeutung* bezeichnet eine Bedeutungsvariante der Komponenten, die keinen Phraseologismus darstellt.¹ Der Satz gilt als obere Grenze phraseologischer Wortverbindungen (z. B. ein Sprichwort). Die Festigkeit einer Wortverbindung kann mit dem Kriterium der *Gebräuchlichkeit* in Verbindung gebracht werden, aber auch mit psycholinguistischer Evidenz, nachgewiesen z. B. durch Tests. Mit *struktureller Festigkeit* bezeichnet Burger die Tatsache, dass phraseologische Wortverbindungen häufig gegenüber den normalen morphosyntaktischen und semantischen Regeln Irregularitäten aufweisen (z. B. bezüglich der Valenz).

Die strukturelle Festigkeit wird durch Variation und Modifikation relativiert. Die *Variation* bezeichnet Varianten von ähnlichen Phraseologismen, z. B. «sich jmdm in den Weg stellen» vs. «jmdm im Weg stehen» vs. «jmdm aus dem Weg gehen», die durch unterschiedliche Aktionsarten gekennzeichnet sind. Demgegenüber bezeichnet *Modifikation* die vom Sprecher in einem bestimmten Kontext hergestellte Abwandlung eines Phraseologismus (Bsp. «Schlechter Rat ist teuer» (Burger, 2007, S. 25–27). Abgesehen davon geschehen bei der Anwendung von Phraseologismen Versprecher und Fehler, und der Übergang von bewusst eingesetzter Modifikation und «fehlerhafter» Verwendung einer Wortverbindung ist fließend. Die *Modifikation* im

¹Beispiel: «sich die Zähne putzen» repräsentiert sowohl die phraseologische als auch die wörtliche Bedeutung (Burger, 2007, S. 13), «die Zähne des Getriebes wurden nie geputzt» repräsentiert aber eine freie Bedeutung, weil hier «putzen» z. B. durch «reinigen» ersetzt werden könnte.

Sinne von Burger kann mit der *exploitation* im Sinne von Hanks (siehe Kap. 6.11) verglichen werden. Die *pragmatische Festigkeit* charakterisiert Routineformeln wie «Guten Tag», die der Gesprächsorganisation dienen und wenig semantischen Gehalt aufweisen (Imo, 2007, S. 24), was teilweise auch für den Bereich der *formulaic language* gilt (siehe Kap. 4.7).

Phraseologie ist eine linguistische Disziplin, deren Grundbegriffe vor allem in der Sowjetunion seit den 1940er Jahren erarbeitet wurden. Aus der russischen Tradition der Phraseologie entstanden *Top-Down-Typologien*, die ein Kontinuum von freien Kombinationen über mehrere Stufen zu fixierteren Kollokationen und zu puren Idiomen als Gegenstand der Phraseologie postulierten (Burger, Dobrovolskij, Kühn & Norrick, 2007; Granger & Paquot, 2008). Demgegenüber steht der von John Sinclair begründete *Bottom-Up-Ansatz*, der auf lexikalischen Kookurrenzen beruht (basierend auf realen Sprachdaten aus Korpora). Die Phraseologie beschäftigt sich mit dem Bereich zwischen Lexikon und Grammatik und grenzt an die Gebiete der Semantik, der Syntax, Morphologie und des Diskurses; eine präzise Abgrenzung gibt es nicht (Granger & Paquot, 2008, S. 30).

Eine geschichtliche Darstellung der Phraseologie findet sich auch bei Feilke (2004, S. 50). Er zeichnet eine Entwicklungslinie der Phraseologie von der vorpragmatischen Phraseologie der 1960er/70er Jahre über die Pragmatisierung der Idiomatik zu einer Pragmatisierung und Kontextualisierung der Phraseologie. Wichtig für das Verständnis von Phraseologismen ist die Kontextualisierung und der *Kontext*. Dieser wird nicht nur durch die Situation vorgegeben, sondern wird von den Beteiligten aktiv (mit)konstruiert (*role taking* und *role making*, vgl. Kap. 3.5). Mit der *Pragmatisierung* der Phraseologie rückt die Frage ins Zentrum der Aufmerksamkeit, «welche Ausdrücke Sprecherinnen und Sprecher pragmatisch als Einheiten der Sprachproduktion und des Sprachverstehens behandeln» (Feilke, 2004, S. 41). Feilke vertritt die Ansicht, dass Phraseologismen eher die Standardeinheiten der Sprache sind als die Ausnahmen. Es zeigt sich damit, dass die theoretisch unbeschränkte Kombinierbarkeit von Wörtern und Ausdrücken, wie sie von den Generativisten postuliert wird, in der praktischen Sprachverwendung nicht eine so grosse Rolle spielt; die Sprachverwendung ist vielmehr stark durch Phraseologismen und feste (idiomatische) Ausdrücke geprägt (vgl. Kap. 4.7). Andererseits sind die festen Ausdrücke auch wieder nicht so fest, denn, wie Korpusuntersuchungen zeigen, sind auch die festen Ausdrücke sehr variabel und flexibel (Moon, 1998).

Granger und Meunier (2008) bezeichnen die Phraseologie als eine Teildisziplin der Lexikologie für das Studium von *Mehrwortausdrücken* (MWE). Die Korpuslinguistik stellt Methoden zur Verfügung, mit denen das Vor-

kommen von MWE in grossen Textsammlungen überprüft werden kann. Damit erfolgt eine Ausdehnung der Phraseologie, die sich ehemals nur mit fixierten und semantisch intransparenten (opaken) Ausdrücken beschäftigte. Phraseologische Fragestellungen werden auch in der Gesprächsforschung, der Korpuslinguistik, in der Konstruktionsgrammatik, in der Psycholinguistik (Spracherwerb, Sprachverstehen und Sprachproduktion) und in weiteren Disziplinen untersucht. Teilweise haben diese Disziplinen bisher voneinander nur wenig Notiz genommen; so besteht die Gefahr, dass die Forscher in verschiedenen Bereichen das Rad immer wieder neu erfinden. Dies ist auch ein grundsätzliches Anliegen des vorliegenden Projektes. Die Beschreibung und Anwendung von MWE soll disziplinübergreifend erfolgen. Immer wieder ergeben sich dadurch aufschlussreiche Querverbindungen und Lösungsansätze, die für die Erzählanalyse und das erzählanalytische Lexikon von grossem Nutzen sind.

Granger und Paquot (2008, S. 42) legen eine ausführliche Typologie von Phraseologismen vor und unterscheiden als Hauptgruppen *referentielle* Phraseologismen von *textuellen* und *kommunikativen*:

1. *Referential phrasemes*: Lexical collocations, idioms, irreversible bi- and trinomials, similes, compounds, grammatical collocations, phrasal verbs.
2. *Textual phrasemes*: Complex prepositions, complex conjunctions, linking adverbials, textual sentence stems.
3. *Communicative phrasemes*: speech act formulae, attitudinal formulae (including attitudinal sentence stems), proverbs and proverb fragments, commonplaces, slogans, idiomatic sentences, quotations.

Heid (2008, S. 340) dagegen führt eine sehr praktisch orientierte, wenn auch heterogene Auflistung von Phraseologismustypen mit verschiedenen Klassifikationskriterien ein, die besonders mit dem Blick auf NLP-Applikationen entworfen wurde:

- MW named entities
- MW function words
- MW adjectives
- Verb and particle constructions
- Nominal compounds

- MW verbal expressions and complex predicates
- Collocations, including light verb constructions
- Stereotyped comparisons: «wie Feuer und Wasser sein»
- Idiomatic expressions
- Binomials: «Salz und Pfeffer»
- Proverbs, quotes and sayings. Beispiel: «Morgenstund hat Gold im Mund»
- Spoken language particle clusters: «gerade mal...»
- Routine formulae of conversation

MWE kommen im Lexikon häufiger vor als Einzelworte. Heid betont die *lexikalische Produktivität*: Eine fest vorgegebene Liste von Phraseologismen deckt nie alle von Sprechern kreativ neu produzierten Ausdrücke ab! Das Lexikon sollte deshalb kombiniert werden mit Regeln, die es ermöglichen, auch ad hoc gebildete MWE-Items zu berücksichtigen (Heid, 2008, S. 349). Das grosse Variationspotenzial und die produktive Erweiterung (exploitation) idiomatischer Ausdrücke sind grosse Herausforderungen einerseits für das Lexikon, aber vor allem auch für das Erkennen der MWE in den Textdaten. Das Feststellen von variablen Idiomanteilen erfordert sehr grosse Korpora, weil diese Phänomene selten sind.

4.2 Mehrwortausdrücke, Multiword Expressions

Im Gegensatz zum Begriff der *Kollokation* ist der Begriff «Mehrwortausdruck» (multi word expression, MWE) weniger durch unterschiedliche Definitionen vorbelastet und hat einen breiteren Geltungsbereich. Mehrwortausdruck wird oft synonym zum Begriff Phraseologismus verwendet. MWE machen einen grossen Teil der Sprachverwendung aus und sind eine grosse Herausforderung für Grammatik und Lexikon. Der grosse Stellenwert von Phraseologismen in der Sprachverwendung zeigt sich z. B. deutlich in den Fehlern von L2-Lernern (Fellbaum, 2007b, S. 2). Mehrwortausdrücke sind teilweise lexikalisch verfestigt (feste Wortverbindungen, Idiome, Metaphern), aber grösstenteils sind sie morphosyntaktisch und lexikalisch mehr oder weniger flexibel.

Der englische Korpuslinguist und Lexikograph John Sinclair erarbeitete wichtige Zusammenhänge zwischen Phraseologie und Lexikographie auf dem Hintergrund umfangreicher Korpusuntersuchungen und bezeichnete Phraseologismen (collocations) als Schlüssel zur Sprachproduktion und zum Sprachverstehen; Bedeutung entsteht durch und in phraseologischen Textmustern (Moon, 2008). Sprecher haben eine grosse Menge von halb-vorgefertigten (semi-constructed) Satzfragmenten zur Verfügung, die in der Kommunikation rasch abgerufen werden können. Diese Phrasen reichen von fixierten MWE bis zu loseren phraseologischen Präferenzen (Moon, 2008, S. 3). Sinclair (1991) unterscheidet zwischen dem *open-choice principle* und dem *idiom principle*. Gemäss open-choice principle produzieren Sprecher Äusserungen aus einzelnen linguistischen Komponenten (Substantiven, Verben, Präpositionen etc.), während das idiom principle besagt, dass Sprecher vorgefertigte Bausteine abrufen. Sinclair zufolge kommt dem idiom principle grössere Bedeutung zu. (Hanks, 2010) betont, dass es wichtig ist, die beiden Prinzipien gleichwertig anzuwenden, wenn die Bedeutung eines Ausdrucks analysiert werden soll:

The *open-choice principle* (alternatively called the terminological tendency) is: a way of seeing language as the result of a very large number of complex choices. At each point where a unit is complete (a word or a phrase or a clause), a large range of choices opens up and the only restraint is grammaticalness.

The *idiom principle* (alternatively called the phraseological tendency), on the other hand, is described in the following terms: Many choices within language have little or nothing to do with the world outside. [...] A language user has available to him or her a large number of semi pre-constructed phrases that constitute single choices. (Sinclair 1991, zit. nach Hanks, 2010, S. 7)

Sinclair beobachtet die Tendenz, dass hochfrequente Wörter weniger unabhängige (literale) Bedeutung haben; d. h. der Kontext wird dann immer wichtiger. Im *Cobuild project*² (1980-1987, vgl. Moon, 2008, S. 5–10) stellte Sinclair die folgenden Statements auf:

- Eine lexikalische Einheit kann von beliebiger Länge sein (d. h. kann aus einem oder mehreren Worten bestehen).
- Es gibt keine Lexikoneinträge mit mehr als einer Bedeutung!

²Collins COBUILD English Language Dictionary (1987, 2nd edition 1995, 3rd edition 2001, 4th edition 2003, 5th edition 2006), basierend auf dem COBUILD-Korpus.

- Die Bedeutungsfestlegung für einen Eintrag ist nicht ohne Kontext möglich!
- Aus diesen Prämissen folgt, dass die Items dieses Wörterbuches Phraseologismen sind; das ultimative Wörterbuch ist ein Kollokations- oder Phraseologismen-Wörterbuch!

Die aufgelisteten Prinzipien und Annahmen gehen stark in Richtung der Annahmen der Konstruktionsgrammatik und werden später in die Anforderungsliste für das JAKOB-Lexikon eingehen.

Phraseologismen verändern das Verständnis von Valenzen. Keil (1997, S. 61) diskutiert in ihrer Arbeit zu verbalen Phraseologismen verschiedene Möglichkeiten. Bezieht sich die Valenz im Beispiel (5) auf das Verb «bringen» (a) oder auf den Phraseologismus «auf die Palme bringen» (b)? Vom phraseologischen Standpunkt aus ist die wendungsexterne Valenz aus Beispiel (b) vorzuziehen, da nur diese dem Phraseologismus gerecht wird.

- (5) Jens bringt Ute auf die Palme
- [bringt (Ute, Jens, auf die Palme)]
 - [bringt auf die Palme (Ute, Jens)]

Entsprechend unterscheidet man *null*, *ein-*, *zwei-* und *dreiwertige Phraseologismen* gemäss Anzahl der «extra-phraseologischen» Komponenten. In diesem Sinne ist «jmd bringt jmdn auf die Palme» ein zweiwertiger Phraseologismus. Ausserdem gibt es *eindeutige und mehrdeutige Valenzmodelle*, die sich darin unterscheiden, ob das für den phraseologischen Ausdruck gültige Valenzmodell auch für die freie Verwendung gilt oder ob es davon abweicht. Ein mehrdeutiges Valenzmodell zeigt sich z. B. in den Ausdrücken «Kreide fressen» vs. «Gras fressen». Hier kann auf Grund des Valenzmodells nicht entschieden werden, ob ein Phraseologismus vorliegt. Ein abweichendes Valenzmodell ist jedoch eindeutig und für das betreffende Idiom abweichend vom nicht-idiomatischen Gebrauch, wie z. B. «jmd hat einen Narren an jmdm gefressen» (Keil, 1997, S. 61). Weitere Beispiele finden sich im *Phraseologismen-Index* (S. 241f) (siehe auch Kap.8.2).

Villavicencio, Bond, Korhonen und McCarthy (2005) besprechen in einer Sonderausgabe der Zeitschrift *Computer Speech & Language* aktuelle Theorien, Definitionen und Klassifikationen zu *multiword expressions* und stellen fest, dass in vielen praktischen Lexikon- und Wörterbuchanwendungen der «words with spaces»-Zugang zu MWE gewählt wird. Demgemäss sind MWE lexikographisch und für die linguistische Verarbeitung wie grosse Wörter zu behandeln. Die meisten NLP-Applikationen ignorieren aber

MWE, behandeln sie als invariante Ausdrücke oder listen die Komponenten einzeln auf, womit die Ausdrücke nicht adäquat beschrieben und analysiert werden können. Für nutzbringende NLP-Applikationen ist es deshalb nötig, potentielle MWE und ihre Variationen zu erkennen und zu interpretieren (S. 366). Neben verschiedenen Definitionen für MWE schlagen die Autorinnen eine einfache und klare Klassifizierung von MWE vor (Villavicencio et al., 2005, S. 367):

1. *Institutionalisierte MWE* (institutionalised phrases, set-phrases): diese sind zwar kompositional, aber konventionalisiert (Bsp. «Zähne putzen» ist die übliche Ausdruckweise, während «Zähne waschen» nicht gebräuchlich ist). Sollen voll kompositionale Phrasen lexikalisiert werden? Für die Sprachrezeption scheint dies nicht nötig zu sein, für die Sprachproduktion, z. B. für L2-Lerner und für Übersetzungssysteme ist es hingegen wichtig, diese konventionalisierten (aber nicht semantisch motivierten) Beschränkungen zu kennen (vgl. Feilke, 1996, S. 164).
2. *Lexikalisierte MWE* (lexicalized phrases) sind semantisch oder syntaktisch in unterschiedlichem Grade idiosynkratisch und werden in folgende Gruppen aufgeteilt:
 - *Syntaktisch variable MWE* (syntactically flexible expressions): Einige MWE können als kompositional angesehen werden, weil sie aus Komponenten bestehen, die variabel zusammengesetzt werden können (z. B. «den Stein ins Rollen bringen» vs. «die Kampagne ins Rollen bringen»). Zu dieser Gruppe gehören auch Verb-Partikel-Konstruktionen (verb particle constructions); die Partikeln können sich in unterschiedlichen Positionen und Abständen vom Verb befinden.
 - *Teilweise fixierte MWE* (semi-fixed expressions) erlauben gewisse lexikalische Variationen (Bsp. «einen *grossen* Bären aufbinden»).
 - Als *feste MWE* (fixed expressions) werden voll lexikalisierte und invariable Ausdrücke bezeichnet.

Aus diesen Definitionen wird noch einmal eine potentielle Eigenschaft von MWE ersichtlich, nämlich die *Produktivität*. Können ad hoc neue Ausdrücke auf der Basis eines existierenden Phraseologismus gebaut werden oder ist der Ausdruck (z. B. idiomatisch) ganz fixiert? NLP-Anwendungen unterscheiden daher an erster Stelle zwischen kompositionalen und nicht-

oder nur teilweise kompositionalen Ausdrücken. Gemäss Heid (2008, S. 343) ist dies auch die Sicht der Leser oder Hörer, für die die wichtigste Frage bei der Verarbeitung von MWE lautet: Wo beginnt der Mehrwortausdruck und wo endet er?

Burger (2007, S. 62) unterscheidet verschiedene *Lesarten* von Phraseologismen und schlägt damit eine praktisch orientierte Typologie vor:

1. Phraseologismen mit nur *einer Lesart*: Dazu gehören einerseits Ausdrücke ohne Idiomatizität (set-phrase, feste Wortverbindung), und andererseits solche mit starker Idiomatizität (Bsp. «Dank sagen», «klipp und klar»).
2. Phraseologismen mit *zwei disjunktiven Lesarten*, z. B. «jmdm einen Korb geben», «das fünfte Rad am Wagen sein». Die wörtliche und die phraseologische (hier idiomatische) Bedeutung erfordern einen andern Kontext.
3. Phraseologismen mit *zwei Lesarten, die simultan realisiert werden* können, z. B. «die Achseln zucken». Achselzucken kann als nonverbale Handlung verstanden werden, und gleichzeitig kann damit «etwas zu verstehen geben» gemeint sein. Der Gebärde entspricht ein Phraseologismus.
4. *Phraseologismen vom gemischten Typ* bestehen aus mehreren Elementen, die teilweise idiomatisch sein können, z. B. mit einer idiomatischen und einer nicht-idiomatischen Komponente. «Einen Streit vom Zaun brechen» (einen Streit beginnen) enthält die idiomatische Komponente «vom Zaun brechen», während «Streit» nicht-idiomatisch ist. Auch Burger (S. 66) plädiert für die Unterscheidung der externen und internen Valenz. Die externe Valenz des Phraseologismus lautet [brechen vom Zaun (Subjekt, Streit)] (vgl. Bsp. 5); die interne Valenz ergibt hier keinen Sinn, da der Ausdruck nicht frei verwendet werden kann.

Zur semantischen Repräsentation von MWE präsentieren Vigliocco, Meteyard, Andrews und Kousta (2009) eine interessante Theorie, die die wesentlichen Bausteine als Grundlage für die Wort- und Ausdrucksbedeutung formuliert. Hauptsächlich sind es zwei Informationstypen, die Wortbedeutung vermitteln: Der erste Informationstyp ist *experiential information*; Informationen dieses Typs stammen einerseits aus unseren sensomotorischen Interaktionen mit der Umwelt, andererseits aus unserer Erfahrung mit inneren Zuständen, speziell Emotionen. Der zweite Informationstyp stammt

aus unserer *Spracherfahrung* und ist aus dem sprachlichen Kontext abgeleitet, in dem das Wort auftritt. Die Autoren zitieren Studien, die ihre Sicht unterstützen, und präsentieren einige Folgerungen aus ihren Erkenntnissen.

4.2.1 Qualities of Multiword Expressions

Eine ausführliche Liste zu Eigenschaften von MWE aus dem Jahre 2003, eine Art Glossar, ist online zugänglich und soll hier zitiert werden:

<http://mwe.stanford.edu/reading-group.html> (siehe Copestake et al., 2002).

Institutionalisation/conventionalisation: process of an expression becoming recognised and accepted as a lexical item, through consistent use over time (necessary but not sufficient condition on MWE-hood).

Lexicogrammatical fixedness: formal rigidity, preferred lexical realisation, restrictions on aspect, mood, voice, etc. (neither necessary nor sufficient condition on MWE-hood).

- lexicogrammatically fixed MWE: kick the bucket, *the bucket was kicked, *slowly kick the bucket
- lexicogrammatically fixed non-MWE: look like, *(to be) looked like, *is looking like

Semantic/pragmatic non-compositionality: there is a mismatch between the semantics/pragmatics of the parts and the whole; includes the case of the component lexical items having specialised meanings within the context of the MWE, not accessible in simplex contexts (not necessary but sufficient).

- idiomatic expression (non-compositional): the expression is semantically opaque and functions as a gestalt (e.g. kick the bucket)
- idiomatically combining expression (idiosyncratically compositional): the lexical parts can be seen to (post hoc) assume components of the semantics of the whole, whereby the sum of the parts equals the whole (e.g. let the cat out of the bag).

Syntactic irregularity: the expression cannot be parsed based on the simplex morphology (parts of speech) of the components (not necessary but sufficient):

- syntactically-irregular MWEs: all of a sudden, the be all and end all of NP
- syntactically regular MWEs: kick the bucket, fly off the handle

Non-identifiability: when first exposed to the expression, the meaning cannot be predicted from its surface form (not necessary but sufficient).

- idiom of decoding (non-identifiable): «misleading lexical clusters» (e.g. kick the bucket, fly off the handle)
- idiom of encoding (identifiable): idiosyncratic lexical combination; note that all idioms of decoding are also idioms of encoding (example strict idioms of encoding – wide awake, plain truth)

Situatedness: the expression is associated with a fixed pragmatic point (neither necessary nor sufficient):

- situated MWEs: good morning, all aboard
- non-situated MWEs: first off, to and fro

Figuration: the expression encodes some metaphor, metonymy, hyperbole, etc, even if the nature thereof is underspecified (neither necessary nor sufficient):

- figurative expressions: bull market, beat around the bush
- non-figurative expressions: first off, to and fro

Proverbiality: the expression is used «to describe—and implicitly, to explain—a recurrent situation of particular social interest ... in virtue of its resemblance or relation to a scenario involving homely, concrete things and relations» (neither necessary nor sufficient)

Informality: the expression is associated with more informal or colloquial registers (neither necessary nor sufficient)

Affect: the expression encodes a certain evaluation of affective stance toward the thing it denotes (neither necessary nor sufficient)

Anomalous collocations: lexicogrammatically marked (syntactically) ill-formed collocations: (at all, by and large)

cranberry collocations: idiosyncratic lexical component – one or more words found only in that collocation (in retrospect, kith and kin)

defective collocations: idiosyncratic meaning component (in effect, foot the bill)

phraseological collocations: semi-productive constructions, occurring in paradigms (in/into/out of action, on show/display)

Formulae: pragmatically marked [Phraseologismen als konventionelle Äußerungsformen zum Vollzug bestimmter Sprechakte, z.B. Schimpfwörter. ML]

simple formulae/sayings: compositional strings with a special discourse function (alive and well, a horse, a horse, my kingdom for a horse)

metaphorical/literal proverbs: (you can't have your cake and eat it, enough is enough), similes (as good as gold)

Metaphors: semantically marked (non-compositional):

- transparent metaphors: (behind someone's back, pack one's bags)
- semi-transparent metaphors: (on an even keel, pecking order)
- opaque metaphors: (bite the bullet, kick the bucket)

Collocations: compositional word co-occurrence of high frequency:

- semantic collocations: co-occurrence preferences/priming effects (jam with FOOD)
- lexico-semantic collocations: collocation paradigms (rancid butter/fat, face the truth/facts/problem)
- syntactic collocations: fully-productive phraseological collocations (too ... to ...)

4.2.2 Zum Begriff «Kollokation»

Der Begriff Kollokation ist nicht klar zu definieren. Er wird von verschiedenen Autoren und in verschiedenen Kontexten verwendet, die dahinter stehenden Konzepte sind aber sehr unterschiedlich. (Burger, 2007, S. 54) schlägt vor, die Bezeichnung *Kollokation* «für den ganzen Bereich der festen

Wortverbindungen, die nicht oder nur schwach idiomatisch sind, zu verwenden». Nach der Definition von Hausmann (2004) besteht eine Kollokation aus zwei Wörtern, aus der *Basis*, einem Einzelwort ohne Kontext, meist einem Nomen, und aus dem *Kollokator*, z. B. dem Verb.³ Aber man ist sich darüber einig, dass Kollokationen auch aus mehr als zwei Komponenten bestehen können (vgl. Feilke, 1996, S. 116).

Die Definitionskriterien für den Begriff *Kollokation* im Kontextualismus (siehe Kap. 5) (Lemnitzer & Zinsmeister, 2006, S. 28) beschränken sich auf das faktische Miteinander-Vorkommen von zwei oder mehr Wörtern/lexikalischen Einheiten. Dafür werden besser die Begriffe *Koakkurrenz* (Gemeinsames Vorkommen zweier Wörter in einem gemeinsamen Kontext) und *Kolligation* (sprachliche Einheiten, verbunden durch syntaktische Kategorien und Relationen) verwendet (Lemnitzer & Zinsmeister, 2006, S. 30).

Viele Autoren plädieren dafür, dass Kollokationen nicht nur durch Kriterien der Häufigkeit des gemeinsamen Auftretens definiert werden sollten, sondern dass auch interne Bedeutungsstrukturen semantischer, syntaktischer oder konnotativer Art eine wichtige Rolle spielen (Feilke, 1996; Granger & Paquot, 2008; Lemnitzer & Zinsmeister, 2006). Einen interessanten Überblick zu Kollokationsdefinitionen (interessant vor allem im Zusammenhang mit seiner Theorie des *lexical priming*) liefert Hoey (2005). Er zitiert die Definitionen von Partington (1998), der eine textuelle, eine statistische und eine psychologische Definition der Kollokation vorschlägt. Die *textuelle Definition* bezeichnet das gemeinsame Vorkommen von zwei oder mehr Worten innerhalb eines kleinen Textfensters als Kollokation. Diese Definition ist nicht sehr nützlich und wird von Hoey verworfen. Die *statistische Definition* bezeichnet die Beziehung zwischen lexikalischen Items im Kontext mit grösserer als Zufallswahrscheinlichkeit als Kollokation. Hoey betont, dass damit nichts über die Bedeutung des Phänomens gesagt ist. Mit der *psychologischen* oder *assoziativen Definition* bezieht sich Partington auf zwei verschiedene Ansätze. Hallyday & Hasan (1976, zit. nach Hoey, 2005) definieren Kollokation als Oberbegriff für die *Kohäsion*, die aus dem gemeinsamen Vorkommen von lexikalischen Items in bestimmten Kontexten entsteht, während Leech (1974, zit. nach Hoey, 2005) postuliert, dass die Kollokationsbedeutung durch die *Assoziationen* entsteht, die ein Wort durch häufige Kontextworte erhält. Leech nimmt also sowohl die statistische als auch die psychologische Realität in seine Definition auf (Hoey, 2005, S. 3–5). Hoey's eigene Definition der Kollokation lautet schliesslich:

³Kollokatoren brauchen Basen, um Bedeutung zu bekommen: Das Verb «decken» bekommt erst mit dem Begriff «Tisch» eine bestimmte Bedeutung (Hausmann, 2004, S. 312).

So our definition of collocation is that it is a psychological association between words (rather than lemmas) up to four words apart and is evidenced by their occurrence together in corpora more often than is explicable in terms of random distribution. (Hoey, 2005, S. 5)

Feilke lehnt Wiederholung im Sinne von Prägung ab. Das Kriterium des frequenten Gebrauchs sagt über die *Qualifizierung der Einheiten als Zeichen* nichts aus; entscheidend ist alleine die Kenntnis des Gebrauchszusammenhangs (Feilke, 2004, S. 52). Rekurrente Kookkurrenz ist ein heuristisch wertvolles Mittel zur Entdeckung von Kollokationen, aber keine Erklärung für die Aneignung von Kollokationswissen.

Die verschiedenen Sichtweisen und Definitionen können nicht nur als eine Definitionsfrage des Begriffs *Kollokation* angesehen werden, sondern sie zeigen konzentriert an dem einen Begriff die vielfältigen Mechanismen der Sprache als «artikulatorischer Operation (Verknüpfung von Zeichen), konzeptualisierender Tätigkeit und sozialer Handlung» (Feilke, 1996, S. 179).⁴ Für die praktischen Aspekte im vorliegenden Projekt wird so verfahren: der Begriff Kollokation existiert zwar als Klassifikationskriterium, wird aber möglichst vermieden und durch eindeutigere Kategorien ersetzt (vgl. Kap. 9.1.2).

4.3 Idiome und Metaphern

Idiomatische Wortverbindungen vermitteln eine Bedeutung, die nicht aus den Bedeutungen der einzelnen Bestandteile erschlossen werden kann. Diese Besonderheit gilt ebenso für die Metapher und für das Phänomen der Metonymie; Metaphorik und Metonymie spielen bei vielen Idiomen eine Rolle; Wortverbindungen können einer oder mehreren der drei Gruppen angehören. Die Grenzen, z. B. zwischen Metapher und Metonymie, sind oft nicht klar bestimmbar. Bei den *Somatismen* steht z. B. ein Körperteil für den ganzen Menschen. Diese Relation wird *Metonymie* genannt und bezeichnet eine «Bedeutungsverschiebung zu einem Aspekt, der in realer Beziehung zur <eigentlichen> Bedeutung steht». Bei der *Metapher* dagegen erfolgt eine Bedeutungsverschiebung in einen neuen Bereich. Die *Synekdoche* bezeichnet eine Teil-Ganzes-Beziehung und wird meistens unter der Metonymie subsumiert (Burger, 2007, S. 86).

⁴als Analogie zur Trias Syntax, Semantik und Pragmatik.

4.3.1 Idiome

Idiome sind Ausdrücke, die lexikalische, morphosyntaktische, syntaktische oder semantische *Idiosynkrasien* oder Variationen aufweisen (Heid, 2008, S. 344). Lexikalische Idiosynkrasie liegt vor, wenn die Semantik eines Ausdrucks nicht transparent ist (z. B. «ins Gras beißen»). Idiomatische Ausdrücke mit variablen Anteilen kommen häufiger vor als erwartet (Fellbaum, 2007b, S. 12); ihre empirische Feststellung erfordert aber grosse Textmengen, weil es sich dabei, selbst in grossen Korpora, um seltene Phänomene handelt. Morphosyntaktische Idiosynkrasien und Variation betreffen z. B. die Verwendung von Singular vs. Plural, von aktiven vs. passiven Formen. Söhn und Römer (2004, S. 161–163) bezeichnen Idiomatizität «als Lexikalisierung von Syntagmen»⁵ mit folgenden Kennzeichen: Anomalien in Morphemen, syntaktische Charakteristika wie Valenzstruktur («jmdn Lügen strafen») und semantische Indizien wie Selektionsrestriktionen (vgl. auch Fellbaum, 2007b, S. 11–12).

Idiome können kreativ (oder auch unwissentlich) modifiziert werden (*idiom breaking*), wie sich auf Grund von Korpusevidenzen unschwer belegen lässt. Die kanonische Form eines Idioms gibt es nicht Söhn (2006, S. 50). Eine «klassische» Sichtweise besagt, dass Idiome feste Einheiten sind und sich in der Grammatik verhalten wie einfache Lexeme (Fellbaum, 2007a). Korpusdaten zeigen aber, dass viele, wenn nicht die meisten, Idiome veränderlich sind; sie können sowohl lexikalische als auch syntaktische Variationen aufweisen. Fellbaum stellt sich gegen eine Sichtweise, die Idiome lediglich als «lange» lexikalische Items bezeichnet (S. 188).

Zu den Idioms gehören auch die *unikalen Elemente* (engl. «cranberry words» oder «bound words») wie «Maulaffen feilhalten» oder «den Ausschlag geben», die eine Komponente besitzen, die in der heutigen Sprache keine freie Bedeutung (mehr) hat.⁶

Kognitive *Theorien zur Idiomatik* werden von Dobrovol'skij (2004) dargestellt; er unterscheidet eine kognitiv-*psychologisch* ausgerichtete Forschungsrichtung, vertreten vor allem durch Gibbs (1994), der Phraseologismen experimentell untersucht, indem Hypothesen über den Ablauf der kognitiven Verarbeitung empirisch getestet werden. Das Ziel ist es, Prozesse der mentalen Verarbeitung von Idioms zu ergründen; das Forschungsinteresse

⁵ «Idiome sind verfestigte Phrasen, die ihre ursprünglichen Syntagma-Eigenschaften in unterschiedlichem Grade verloren haben und zu Lexikoneinheiten geworden sind.» (Söhn & Römer, 2004, S. 161)

⁶ Im Internet existiert eine kommentierte Liste unikalener Elemente der deutschen Sprache (The Collection of Distributionally Idiosyncratic Items, CoDII), URL: <http://www.sfb441.uni-tuebingen.de/a5/codii/>

liegt eher im Bereich der Psycholinguistik. Die kognitiv-*linguistische* Richtung wird von Dobrovol'skij selbst vertreten; er untersucht im Rahmen des kognitiv-linguistischen Paradigmas mit Hilfe von Heuristiken die sprachlichen Besonderheiten der Idiomatik (Dobrovol'skij, 2004, S. 118). So wirft er die Frage auf, ob «ins Gras beißen» und «den Löffel abgeben» Synonyme sind. Im Sinne von Dobrovol'skij sind sie dies nicht (obwohl sie in ihrer Kernbedeutung übereinstimmen), weil sie sich in ihren bildlichen Komponenten unterscheiden (S. 120). Dieser Befund ist intuitiv nachvollziehbar. Die gleiche Überlegung wird für die dramaturgische Kodierung in der Erzählanalyse gemacht, wenn Phraseologismen mit einem Doppelcode ausgezeichnet werden, der einerseits die semantische Kernbedeutung abbildet und andererseits auch auf den Quellbereich des Ausdrucks Bezug nimmt. In diesem Sinne würden im JAKOB-Lexikon die beiden erwähnten Idiome als «near-synonym» betrachtet.

(6) an allen Ecken und Enden

(7) auf Schritt und Tritt

Für die Idiome in den Beispielen (6) und (7) stellt Dobrovol'skij (2004) die These auf, dass (7) sich nur auf bebaute und bewohnte Räume beziehen kann, während (6) an vielen Stellen geschehen kann. Dieser Unterschied lässt sich allerdings in Korpusdaten nicht nachweisen, was zeigt, dass die Heuristiken der kognitiven Semantik zwar «Auskunft [geben] über die semantischen Potenzen eines Konzepts, nicht aber über ihre de-facto-Realisierungen.» (S. 121–123).

4.3.2 Metaphern

Zum Gebrauch von Metaphern gibt es viele Theorien und Konzepte. In ihrer Übersichtsarbeit beschreibt Deignan (2005) die Entwicklung der Metaphertheorien ausgehend von den älteren Theorien, die die Metapher als Verzierung, als dekorative Variante der Sprache, oder negativ ausgedrückt als unnötige Ablenkung vom Wesentlichen betrachteten (sie bezeichnet diese Tradition als «decorative view of metaphor») zu den kognitiven (oder konzeptuellen) Metaphertheorien, wie sie z. B. von Lakoff und Johnson (1980) und Lakoff (1987) vertreten werden, bis zu neueren Metaphernkonzepten, die sich an der Korpuslinguistik orientieren und ihre Erkenntnisse auf Korpusevidenzen stützen.

Um gewisse Dinge zu sagen oder zu kommunizieren, haben wir gar keine andere Wahl, als Metaphern zu verwenden. «Why do we use expressions metaphorically instead of saying exactly and literally what we mean?» (Searle,

1993, zit. nach Deignan, 2005, S. 17). Deignan braucht den Ausdruck «der *Lauf* des Lebens» (the *course* of human life) als Beispiel, um zu zeigen, dass es oft nicht möglich ist einen Sachverhalt «exakt und wörtlich» zu beschreiben.

Als wichtige Annahmen einer kognitiven Metapherntheorie zählt Deignan (2005, S. 13) die folgenden Leitsätze auf:

- Metaphern strukturieren das Denken.
- Metaphern strukturieren das Wissen.
- Die Metapher ist grundlegend für das abstrakte Sprechen.
- Metaphern gründen auf körperlicher Erfahrung (physical experience).
- Die Metapher ist ideologisch.

Es ist wichtig, das metaphorische *Denken* vom metaphorischen *Sprechen* zu unterscheiden. Das Denken entspricht der *konzeptuellen*, das Sprechen der *linguistischen* Metapher. Die konzeptuelle Metapher ist ein gedankliches Konstrukt (Schema), während die linguistische Metapher die Realisierung eines Metaphernkonzeptes ist, die konkrete Instanzierung, wie sie im Sprachgebrauch umgesetzt wird (Deignan, 2005, S. 14).

- (8) a. Metapher: Dampf ablassen
 b. Konzept: ÄRGER ALS ERHITZTE FLÜSSIGKEIT IN EINEM BEHÄLT-
 TER

Eine *konzeptuelle Metapher* verbindet zwei semantische Felder (domains), in der Regel ein konkretes und ein abstraktes (source und target domain), bei Burger (2007) *Ausgangsbereich* und *Zielbereich* genannt. Die Metapher «Dampf ablassen» ist eine konventionelle Metapher, die lexikalisiert ist, und deren Bedeutung nicht aus Situation und Kontext erschlossen werden muss.⁷ Neue Metaphern können aber auch ad hoc kreiert werden. So ist die metaphorische Konstruktion «Dampf aufbauen» eine kreative Neuformulierung (exploitation), die das Konzept hinter der Konstruktion «Dampf ablassen» antonymisch verwendet (unkonventionelle Metapher).⁸

⁷Dies stimmt nicht ganz; es ist ein Unterschied, ob die Szene in einem Dampflok-Depot spielt oder ob über die psychische Befindlichkeit einer Person berichtet wird.

⁸Beide Ausdrücke kommen im selben Gesprächskorpus vor; die Unterscheidung wird durch die Häufigkeit des Auftretens der beiden Formulierungen nahe gelegt (siehe Kap. 9.2.5).

In den Kapiteln zur Narrativik und Gesprächsforschung zeigen mehrere Autoren den Nutzen der Metaphernverwendung in psychotherapeutischen Gesprächen auf. Shadrach Long und Lepper (2008) weisen nach, dass im Therapiegespräch *neue* und unkonventionelle Metaphern vor allem in ausgedehnten Redephasen, in Momenten der tieferen Einsicht in Zusammenhänge, wenn ein Problem gelöst wird oder beim Entwickeln von wichtigen Themen kreiert werden (S. 345). Die Verwendung von Metaphern im Therapiegespräch ist für die Beteiligten wichtig und nützlich, weil sie Mentalisierungsprozesse ermöglicht und erleichtert.⁹ Unter dem Begriff «Metapher» fassen die Autorinnen allerdings in diesem Zusammenhang ein breites Spektrum von Phraseologismen (figurative language) zusammen: Metapher, Vergleich (simile), Metonymie, Hyperbel (Übertreibung), Untertreibung (understatement), Oxymoron (Gegensatz, Bsp. «alter Knabe»), Synekdoche, Idiom, Ironie (Shadrach Long & Lepper, 2008, S. 348).

Der *Vergleich* wird bei Burger als *komparativer Phraseologismus* bezeichnet und besteht aus drei Hauptelementen (am Beispiel von «jemand ist flink wie ein Wiesel»): 1) primum oder comparandum, das Vergleichsobjekt, der Vergleichsempfänger (jemand). 2) tertium comparationis (flink). 3) secundum oder comparatum, das Wiesel als Vergleichsmass oder Vergleichssponder (Burger, 2007, S. 46). Im Gegensatz zur Metapher, die auch einen Vergleich impliziert wird im komparativen Phraseologismus der Vergleich explizit mit dem Wort «wie» (Vergleichssignal) durchgeführt.

Zum Phänomen des *Vergleichs* analysiert Hanks (2005) die englische Präposition «like», basierend auf deren Vorkommen im British National Corpus (BNC), wie z. B. «look like a boiled frog». Der Vergleich wird vor allem mit Verben der Wahrnehmung verwendet, aber auch um eine unscharfe Menge (fuzzy set) anzugeben: «Leute wie Ärzte und Krankenschwestern». Um etwas Neues auszudrücken, bedient man sich gerne bekannter Konzepte, d. h. dass die beiden zu vergleichenden Begriffe nicht der gleichen Begriffskategorie (lexical set) angehören dürfen (da sonst nicht Unbekanntes mit Bekanntem verglichen werden kann). Diese Annahme wird jedoch im (englischen) Sprachgebrauch nicht unterstützt. Die weitaus am häufigsten in den Vergleichen verwendeten Nomen gehören den semantischen Typen «Human» (in verschiedenen Rollen) oder «Animal» an. Fazit: Jedes häufige Gleichnis-Nomen hat ein salientes Attribut, und dieses ist es, was in der Kreation des Vergleichs verwertet wird (Hanks, 2005, S. 8).

⁹Bei der Metaphernanalyse von Gesprächen gibt es aber immer die keineswegs einfache Frage zu berücksichtigen, wie Metaphern und *figuratives Sprechen* erstens definiert und zweitens identifiziert werden sollen. Eine Operationalisierung und Systematisierung der Metaphernanalyse ist deshalb wünschenswert (Schmitt, 2010, in press).

- (9) He later complained he felt like a Spanish general without a horse.

Die Verwendung dieser metaphorischen Aussage evoziert kulturelle Stereotypen, die dem Leser/Hörer aber nicht genau bekannt sein müssen; dieser macht sich seinen Reim aufgrund seiner stereotypen Kenntnisse über Generäle, Kavallerie und Spanier. Die Zielsetzung des Vergleichs ist nicht kognitive Präzision, sondern Vagheit. Vagheit eröffnet kreativ ein Spektrum von Möglichkeiten im Kopf des Hörers, während Präzision einschränkt (Hanks, 2005, S. 10). Auch Feilke (1996, S. 62) weist auf Unterspezifikation und systematische Vagheit von (idiomatischen, metaphorischen oder metonymischen) Ausdrücken in der Sprachproduktion und -rezeption hin. Sprache als Kommunikationsinstrument erfordert Vagheit; die semantische Plastizität von Sprache ist eine notwendige Bedingung der Kommunikation.

Sprechen [heisst] bekanntlich nicht, den Gedanken sprachlich auszudrücken, sondern es geht in der Sprechfähigkeit um den *sprachlichen* Entwurf des Gedankens selbst mittels eines dafür geeigneten Ausdrucks. (Feilke, 1996, S. 112)

Vergleiche nutzen wie Metaphern kulturelle Stereotypen und allgemeines Welt- und Kulturwissen, um Bedeutung zu konstruieren. Metaphern beziehen sich auf ganze Konzepte, Vergleiche fokussieren auf eine saliente Eigenschaft des Vergleichsspenders (vgl. Hanks, 2005, S. 13).

4.3.3 Metonymie

In der kognitiven Linguistik ist die *Metapher* meistens definiert als «kognitives Mapping» zwischen verschiedenen und unabhängigen *Domänen*; ein Phänomen aus einer bestimmten Domäne wird durch eine Analogie aus einer andern Domäne erklärt (Bsp: die konzeptuelle Metapher ARGUMENT IS WAR erklärt die Vorgänge beim Argumentieren mit den Vorgängen beim Krieg). Die *Metonymie* dagegen beruht auf einem Mapping innerhalb derselben kognitiven Domäne (Lakoff & Johnson, 1980), auf einem Bedeutungskontinuum.

Metonymy is a cognitive process in which one conceptual entity, the vehicle, provides mental access to another conceptual entity, the target, within the same idealized cognitive model. (Radden & Kövecses, 1999, zit. nach Panther & Radden, 2005, S. 3)

Domäne (domain) steht hier stellvertretend für Konzepte, die auch als *semantisches Feld* (semantic field) bezeichnet werden, als *idealisiertes kognitives Modell* (idealized cognitive model, ICM) oder in der Framesemantik

als *Frame*. Jedes Frameelement evoziert den Frame als Ganzes und gleichzeitig weitere Frameelemente aus dem verwandten Frame-Netzwerk. Und weil diese Elemente alle nahe beieinander sind, können sie metonymisch verwendet werden (Panther & Radden, 2005, S. 3).

Die Begriffe Metapher und Metonymie überschneiden sich teilweise, psycholinguistisch gesehen besteht jedoch ein Unterschied, ein anderes Processing: Metonymie beinhaltet einen ursächlichen, räumlichen oder zeitlichen Zusammenhang zwischen dem bildlichen Ausdruck und dem eigentlich Gemeinten. Das Metonym «steht für etwas», ist aber im Gegensatz zur Synekdoche kein direkter (physischer) Teil des Gemeinten (Deignan, 2005, S. 59f.). Es gibt allerdings auch Stimmen innerhalb der kognitiven Linguistik, die die scharfe Unterscheidung von Metonymie und Metapher ablehnen (Panther & Radden, 2005, S. 4).

Metonymie wird von Gibbs (1994) so charakterisiert, dass ein gut zugänglicher und verstehbarer Aspekt einer Sache für die Sache als Ganzes steht,¹⁰ wie in den Beispielen (10) und (11).

(10) Das Weisse Haus ...

(11) Die Bank senkt die Zinsen.

Warren (Warren 1999, zitiert nach Deignan, 2005, S. 56) beschreibt als Eigenart der Metonymie, dass der Hörer die *Relation* zwischen wörtlicher und figurativer Bedeutung herausfinden muss. Bei der Metapher hingegen geht es nicht um diese Relation, sondern um die *gemeinsamen Attribute* (shared attributes), die Quell- und Zielbereich verbinden. Gibbs (1994) erläutert anhand des Verstehens von Gedichten die kognitiven und kommunikativen Funktionen der Metonymie, die grundlegende Denk- und Kommunikationsschemata des Menschen darstellen. Wir verstehen ein Gedicht, weil eine Situation evoziert wird. Verstehen hängt von unserer Fähigkeit ab, metonymisch aufgrund von Details einer Szene etwas über die ganze Szene zu folgern. Metonymie ist tief verankert in der menschlichen Kognition, und metonymisches Denken ist ein konzeptuell und pragmatisch allgegenwärtiges Phänomen (Gibbs, 1994).

Es gibt Metonyme, die nur aus dem Kontext und in der Situation verständlich sind (kreativer Sprachgebrauch); aber es gibt ebenso metonymische Ausdrücke, die lexikalisiert und allgemeinverständlich sind. Deignan (2005, S. 55) erwähnt die englischen Ausdrücke «stage» und «screen», die

¹⁰ «People take one well-understood or easily perceived aspect of something to represent or stand for the thing as a whole» (Gibbs, 1994, S. 320).

metonymisch für Bühnenwelt und Filmwelt stehen (Beispiel: «die Pfauenbühne» als Bezeichnung für das Schauspielhaus in Zürich repräsentiert das metonymische Schema RAUM FÜR THEATERINSTITUTION). Die Beispiele (12) und (13) repräsentieren das Schema INSTITUTION FOR PEOPLE.

- (12) Die Schule verlangt viel von den Kindern.
- (13) Die Universität reglementiert den Zugang zum Masterstudium.

Die nicht-konventionelle Metonymie dagegen ist innovativ, neu, nicht lexikalisiert und nur im Kontext verständlich.

- (14) «The ham sandwich is sitting on table 20» (Nunberg, 1979).

Deignan (2005, S. 57) zitiert das berühmte Beispiel von Nunberg (14). Für das Verstehen ist hier Weltwissen erforderlich und nicht lexikalisches oder linguistisches Wissen, d. h. hier versagt das Lexikon, denn das Metonym ist ein temporäres Label für einen bestimmten Referenten. Beispiele von metonymischen Schemata zeigen weitere Besonderheiten:

- (15) PART-FOR-WHOLE (Beispiel: «ein hübsches Gesicht» für eine hübsche Person).
- (16) SALIENT BODY PART FOR PERSON (Beispiele: Der Dickwanst, das Fettgesicht).
- (17) auch BODY PART FOR PERSON (Beispiel: Wir brauchen junge Köpfe an der Uni).
- (18) CONTAINER FOR CONTENT (Beispiel: Jonathan ist im Telefonbuch).

Die Beispiele stammen von Panther und Radden (2005, S. 3). Der Ausdruck *Dickwanst* im Beispiel (16) hat zwei mögliche lexikalische Bedeutungen: «Dicker Bauch» oder «Dicker Mensch» («potbelly has two entrenched lexical senses: large round stomach and person with large round stomach»)¹¹. Nicht alle Elemente eines konzeptuellen Frames eignen sich als metonymische Quellen oder Ziele. Das Beispiel (17) ist verständlich, während der Ausdruck «wir brauchen einige junge Knie» nicht passend für dieses Konzept wäre (nach Panther & Radden, 2005, S. 11).

Was sind die *Diskursfunktionen* von Metonymie? Warum brauchen Sprecher metonymische Ausdrücke als Mittel der Bedeutungsherstellung? Auf den ersten Blick könnte man meinen, dass durch die Suche eines geeigneten Metonyms der kognitive Aufwand steigt. Aber Panther und Radden

¹¹ Übersetzung *potbelly* im Internet: dicker Bauch; bei canoo: dicker Mensch.

(2005, S. 10) argumentieren, dass durch die metonymische Ausdrucksweise der Situation besser angepasste Konzepte abgerufen werden können, die den erhöhten Prozessieraufwand rechtfertigen. Im Falle von konventionalisierten metonymischen Ausdrücken ist es aber auch möglich, den Prozessieraufwand zu minimieren, besonders in der schnell ablaufenden Routine-Kommunikation.

Folgendes Beispiel eines *Verbpatterns*¹² illustriert gleichzeitig mögliche wörtliche, metaphorische und metonymische Bedeutungen; es stammt von Hanks und Pustejovsky (2005, S. 6):

(19) [[Horse | Person]] gallop [Adv[direction]]

Dieses Verbpattern ist in der Realisierungsform mit «Pferd» nicht mehrdeutig («das Pferd galoppiert»); das Pattern ist aber mehrdeutig, weil es metonymisch gemeint sein kann («der Reiter (auf dem Pferderücken) galoppiert»), oder aber metaphorisch, wenn eine Person galoppiert (zu Fuss).¹³ In einigen Fällen kann das (fakultative) Richtungsadverb aus dem Verbpattern Hinweise zur richtigen Interpretation geben.

4.4 Phraseologie und Computerlinguistik

In der Computerlinguistik ist der Begriff der MWE sehr weit gefasst, weiter als im engeren phraseologischen Rahmen und wird so charakterisiert: «the very broad definition of multiword units often used in computational linguistics, which is more practically grounded than based on phraseological theories» (Heid, 2007, S. 1036). Und Rothkegel (2007) beschreibt das Forschungsinteresse der Computerlinguisten an phraseologischen Fragestellungen wie folgt:

Die Basisfrage ist, ob und aufgrund welcher Bedingungen eine Wortkette - sei sie kontinuierlich oder diskontinuierlich im laufenden Text angeordnet - eine lexikalisch-semantische Einheit darstellt und durch welche Art Verfahren dies festgestellt werden kann. Vor allem Letzteres ist das grosse Problem in der automatischen Sprachverarbeitung. (Rothkegel, 2007, S. 1027)

¹²Ein Verbpattern (Hanks, 2008a) ist ein Satzmuster, das die variablen Argumente eines Verbs als semantische Typen anzeigt (siehe Kap. 6.11).

¹³Ein Beispiel aus dem DWDS zeigt, dass auch weitere semantische Typen galoppieren können: «Doch selbst nach diesem Sonntag setzt sich noch nichts in Bewegung. Die Krise galoppiert nicht, sie schleicht» (DIE ZEIT 12.03.1982, S. 2).

Am Beispiel des Phraseologismus «einen Bären aufbinden» zeigt Rothkegel, wie in einer früheren, syntaxzentrierten Phase der maschinellen Verarbeitung von MWE vor allem die Monolexikalisierung das Ziel war: Ausdrücke wie «einen Bären aufbinden» sollten wenn irgend möglich mit einem einfachen Wort wie «belügen» gleichgesetzt werden können. Unter Berücksichtigung der Semantik ergibt sich aber das Problem, dass im Ausdruck «einen Bären aufbinden» die Konstituente Bär durch Attribute erweiterbar ist (z. B. einen «grossen» oder einen «schönen» Bären...), was bei einer Reduktion auf «belügen» nicht berücksichtigt werden könnte (Rothkegel, 2007, S. 1033). Die bildliche und die wörtliche Bedeutung sind aber für die Bedeutungskonstruktion im vorliegenden Ausdruck wichtig; die Reduktion auf ein einfaches Verb würde den oben angesprochenen kreativen Umgang mit Bedeutungspotenzialen ignorieren. Die Erweiterungs- und Modifikationsmöglichkeiten in der Sprachproduktion sind eine grosse Herausforderung bei der Konstruktion von maschinenlesbaren Mehrwort-Lexika

Heid (2007) beschreibt im Phraseologie-Handbuch von Burger et al. (2007) den aktuellen Stand und die Desiderate einer computergestützten (computerlinguistischen) Phraseologie. Es besteht aktuell ein grosser Bedarf an phraseologischen Wörterbüchern in einem standardisierten und für verschiedene Applikationstypen verwendbaren Format. Die aktuell am meisten verwendete und zukunftsorientierte Technik zur Lexikonerstellung ist die Verwendung statistischer Methoden für die korpusbasierte Extraktion von Phraseologismen. Die Resultate sind jedoch nicht immer befriedigend; statistische Techniken sollten deshalb mit Techniken der deskriptiven Linguistik kombiniert werden (Heid, 2007, S. 1036).

Die Fragestellungen und Aufgaben der Computerlinguistik in der Phraseologie können wie folgt umrissen werden:

1. *Identifikation* von Phraseologismen: MWE in elektronischen Texten erkennen.
2. Automatische *Extraktion* von Phraseologismen aus Textkorpora.
3. *Semantische Repräsentation* von Phraseologismen (z. B. auf Grund eines klassenbasierten Ansatzes oder von Verbpattern).
4. *Standardisierung*: MWE-Lexika sollten in standardisierten Formaten bereitgestellt werden.

Heid gibt eine detaillierte Auflistung der Mehrwort-Phänomene, wie sie von der Computerlinguistik untersucht werden. Diese Liste unterscheidet sich von allgemeinen phraseologischen und linguistischen Typologien und

ist auf die speziellen Bedürfnisse der Computerlinguisten zugeschnitten, wird dadurch aber sehr formal und ist relativ weit entfernt von einer kommunikativ und pragmatisch verankerten Sicht auf die Sprachverwendung (Heid, 2007, S. 1038) (vgl. die Liste 4.1, S. 74).

Identifikation und Extraktion von MWE kann im einfachsten Fall so geschehen, dass im Text nach einer oder möglichst vielen nahe beieinander liegenden Komponenten des Ausdrucks gesucht wird. Wesentlich bessere Resultate sind aber zu erreichen, wenn a) die morphosyntaktische Struktur der zu analysierenden Sätze erkannt werden kann (Parsing), und wenn b) als Gegenstück dazu eine detaillierte Repräsentation des gesuchten Phraseologismus im elektronischen Lexikon zur Verfügung steht (vgl. Heid, 2007). *Phraseo-Lex* (Keil, 1997) ist ein Beispiel einer Lexikondatenbank für verbale Phraseologismen (siehe Kap. 7.2.6). Weitere Beispiele sind *PhraseManager* (Pedrazzini, 1994; Tschichold, 2000, 2008), eine Lexikondatenbank, die einen klassenbasierten Ansatz der Darstellung von MWE verfolgt, und die Liste der *idiomatisierten verbalen Phraseologismen* ohne freie Lesart (IVP, Söhn, 2006, siehe Kap. 7.2.5). Im vorliegenden Projekt JAKOB-Lexikon und in der dazu gehörenden Analyseapplikation verfolgen wir den oben erwähnten Weg: Im Lexikon sollen möglichst viele linguistische und soweit möglich auch extra-linguistische Eigenschaften für Phraseologismen und Verbpatterns zur Verfügung gestellt werden; die zu analysierenden Texte werden segmentiert (satzweise) erfasst, ein robuster Parser erkennt die morphosyntaktische Struktur und vergleicht sie mit den Patterns im Lexikon. Dieses Verfahren eignet sich allerdings nur für kurze Texte, die entsprechend vorbereitet (Segmentierung) und nachbehandelt werden müssen.

Die empirische Suche und automatische Extraktion von MWE in Korpora wird oft so vorgenommen, dass zuerst eine syntaxbasierte Extraktion von potentiell zusammen gehörenden Wortpaaren durchgeführt wird und anschliessend diese Wortpaare gemäss Kookkurrenzkriterien sortiert werden (Heid, 2007, S. 1042). Die statistische Analyse der Kookkurrenz von Wortpaaren wurde 1990 eingeführt durch die Erhebung von Wort-Assoziationsmassen (*Mutual Information*) direkt aus elektronischen Korpora (Church & Hanks, 1990).¹⁴ In der Folge wurden verschiedene Konzepte mit unterschiedlichen statistischen Prozeduren entwickelt, um das Kollokationsverhalten zu messen.

Der Nachteil dieser statistischen Methoden ist, dass sie auch Wortpaare finden, die voll kompositional sind, also keine MWE im Sinne der Phraseologie, und, wie oben bei der Beschreibung der Kollokation erwähnt, reicht es

¹⁴vgl. Kapitel (5.1).

nicht, nur Wortpaare zu finden; Phraseologismen bestehen zum grossen Teil aus grösseren Einheiten. Ausserdem geht es nicht nur um Wortpaare und Wortgruppen, sondern z. B. auch um kontextabhängige morphosyntaktische und semantische Eigenschaften, und um mögliche Variationen innerhalb der Wortgruppen, so dass die statistischen Erkenntnisse mit qualitativen linguistischen Methoden ergänzt werden müssen.

4.4.1 Word Sense Disambiguation (WSD)

Unter diesem Begriff existiert eine eigene Forschungsrichtung innerhalb der Computerlinguistik, die sich mit der Frage beschäftigt, wie der Wortsinn im Kontext gefunden und verschiedene Lesarten disambiguiert werden können; eine Übersicht über den aktuellen Forschungsstand geben Agirre und Edmonds (2006b). WSD ist eine NLP-Aufgabe (natural language processing). Die Forscher in diesem Spezialgebiet gehen ebenfalls davon aus, dass Wörter aus Prinzip mehrdeutig sind; paradoxerweise gibt es aber für einen realen Leser/Hörer in einem realen Text/Gespräch kaum eine wirkliche Ambiguität (Agirre & Edmonds, 2006a)! WSD bezieht sich auf eine Anwendung der Computerlinguistik, entweder als generische Anwendung oder integriert in eine NLP-Applikation. Verschiedene methodische und technische Zugänge werden verwendet, so ein wissensbasiertes Vorgehen, aber auch statistische Verfahren mit oder ohne Supervision.¹⁵

Disambiguierung des Wortsinns ist ein Anliegen in vielen der hier besprochenen Verfahren und Methoden. Speziell auch die *Corpus Pattern Analysis* (CPA) beschäftigt sich damit als zentralem Anliegen, wenn die entsprechenden Studien sich auch nicht ausdrücklich auf WSD beziehen. Bedeutung wird eher dem Kontext zugeschrieben als einzelnen Wörtern. Deshalb werden z. B. im PDEV (pattern dictionary of english verbs, siehe Kap. 6.11) nicht einzelne Verbbedeutungen aufgelistet; jedes Verb ist stattdessen assoziiert mit einem oder meistens mehreren syntagmatischen Mustern, die den selektionalen Kontext bilden (Pustejovsky, Hanks & Rumshisky, 2004, S. 1). Zum Kontext gehören thematische und lokale Eigenschaften, die sich in einem gewissen Textausschnitt um ein Zielwort herum befinden; die lokalen Features werden bestimmt durch N-grams, die mit dem Zielwort assoziiert sind (Pustejovsky et al., 2004, S. 1). Diese Gedanken sind teilweise (oder analog) im Konzept der <topic-queue> in der JAKOB-Applikation umgesetzt: Der thematische Kontext eines Ausdrucks aus dem Lexikon wird für

¹⁵Basic approaches: knowledge-based; unsupervised and corpus-based; supervised and corpus-based; combinations (Agirre & Edmonds, 2006a, S. 13).

den unmittelbar folgenden Text als lokale thematische Vorgabe für die Disambiguierung verwendet.¹⁶

Kritische Stimmen betrachten die Bemühungen der WSD als gescheitert, wie Hanks erklärt:

Traditionally, the methodology of word sense disambiguation (WSD) takes an inventory of word senses from a dictionary such as LDOCE (Longman Dictionary of Contemporary English) and then seeks contextual criteria for the realization of each sense. It is now acknowledged by some of the leading practitioners in the field (e.g. Ide and Wilks 2006) that this procedure has been a failure. One reason for this failure is excessive reliance on reductionist theories of semantic compositionality. Another is failure to pay sufficient attention to the phenomenon of collocation. Another is an expectation of certainty, for in matters of meaning only probabilities may be expected. (Hanks, 2010, S. 9)

4.5 Phraseologie und Psycholinguistik

Häcki Buhofer (2004) stellt interessante Überlegungen zur Verwendung und zum Verstehen von Phraseologismen aus psycholinguistischer Sicht an. So kann sich in der sprachlichen Verständigung eventuell eine *Doppelkodierung* von Gesagtem und Gemeintem ergeben. Die Bedeutung, die im Gespräch bei der Sprachproduktion durch einen Phraseologismus angestrebt wird, wird bei der Rezeption unter Umständen nicht bemerkt oder anders interpretiert (die Autorin erwähnt als Beispiel Werbetexte, die bewusst ein Sprachspiel inszenieren). Diesem Umstand muss in der Gesprächsanalyse Rechnung getragen werden.

Eine <ordentliche> semiotische Herleitung der Bedeutung von Phraseologismen, die wörtliche und übertragene Bedeutung sauber trennt und oft als *logisch* bezeichnet wird, kommt den Bedürfnissen der Analyse entgegen, ist aber nicht treffend für die Vorgänge, die beim Sprachgebrauch bzw. beim Sprachverstehen ablaufen. (Häcki Buhofer, 2004, S. 146)

Die sprachliche Form des Dargebotenen wird vom Hörer in der Regel vergessen, wenn er verstanden hat. Die phraseologische Form bleibt jedoch

¹⁶Dieses Features ist jedoch noch im Aufbau und erst bei wenigen Lexikoneinträgen verwirklicht.

stärker im Gedächtnis erhalten als diejenige einfacher Wörter, auch weil der Verstehensprozess eventuell länger dauert. Dazu gibt es eine neurologische Interpretation: Phraseologismen bilden Knoten, deren Verbindungen besonders leicht aktiviert werden können (Häcki Buhofer, 2004, S. 147). Verstehen kann wie Sprechen als Konstruktionsvorgang betrachtet werden. Bedeutung wird konstruiert; verstanden hat man, wenn die Bedeutung der rezipierten Äusserung als ausreichend akzeptiert wird (S. 150).

Häcki Buhofer (2004, S. 161) bemängelt, dass in der Linguistik oft rekonstruiert und modelliert wird und dass dadurch Produktions- und Verstehensprozesse von Phraseologismen als spezielle Leistungen der Sprachkompetenz erscheinen. Sie hält dagegen, dass die phraseologischen Ausdrucksmöglichkeiten zum normalen Alltagsgebrauch gehören und psychologische Realität abbilden. Besonderheiten und «Fehler» bei der Rezeption von Phraseologismen müssen im Kontext *produktiven Verstehens* gesehen werden.

Eine ältere Sicht auf Phraseologismen (von Gibbs «standard pragmatic view» genannt) behauptet, dass Hörer zuerst die *wörtliche Bedeutung* (literal meaning) eines Ausdrucks verarbeiten müssen und erst dann unter Zuhilfenahme von pragmatischen Kontextinformationen auch *figurative Bedeutungen* analysieren. Dies setzt einen zeitraubenden Übersetzungsvorgang voraus, der nicht plausibel ist und von psycholinguistischen Forschungsergebnissen grösstenteils widerlegt wurde (z. B. Gibbs, 2002). Diese Sicht wird auch von Feilke (1996) geteilt:

Nach diesem Modell würde der Hörer nach einem Vergleich von Kontext und literaler Äusserungsbedeutung zu der Feststellung kommen: Hier ist nicht die wörtliche Bedeutung gemeint! und dann in einem zweiten Schritt vor dem Hintergrund des aktuellen Kontextes die metaphorische Bedeutung konstruieren. Nach allen bisherigen Ergebnissen psycholinguistischer Forschung gibt es dafür jedoch keine Anhaltspunkte. (S. 121)

Im Gegensatz dazu steht die von Gibbs *direct access view* genannte Sicht, wonach nonliterale oder figurative Bedeutung verstanden wird ohne vorherige Analyse der wörtlichen Bedeutung, vorausgesetzt, dass genügend Kontextinformationen zur Verfügung stehen (Gibbs, 2002). Wie Feilke (1996) betont,

haben wir es in der Sprache *immer mit den Funktions- und Gebrauchsbedeutungen* und nicht mehr mit den Motivationsbedeutungen ursprünglich einmal zum Beispiel metaphorisch motivierter Ausdrücke zu tun. (S. 121)

Er zeigt dies am Beispiel des Phraseologismus «im Regen stehen», dessen Bedeutungserschliessung nicht Sachwissen erfordert, sondern Kenntnis der Funktion und des Gebrauchs, also Erfahrung im Umgang mit diesem Ausdruck (S. 121). Diese kann allerdings beim Verstehen neuer Metaphern nicht vorausgesetzt werden. Aber auch hier ist nicht der Rekurs auf die wörtliche Bedeutung zentral, sondern eher der Rückgriff auf unsere ganze Sprachbiographie, die sprachliche Ausdrücke mit Situations- und Kommunikationskontexten verknüpft.¹⁷

Für eine praktische Lexikonumsetzung sehr wichtige Überlegungen zur Bedeutungskonstruktion regt Feilke mit seiner Darstellung des einfachen Lexems «Lage» an. Lage ist semantisch merkmalslos und gewinnt erst im Kontext (Prägungskontext) eine Bedeutung (S. 129):

- (20) a. die Lage der Nation
- b. die attraktive Lage
- c. die (an)gespannte Lage
- d. die Seitenlage
- e. mittlere Lagen

Gemäss Feilke gibt es für die *lexikalische Einordnung* dieser Ausdrücke zwei Möglichkeiten: Einerseits den *semasiologischen Ansatz*, der ein abstraktes Lexem «Lage» postuliert und die jeweils aktuelle Bedeutung über kontextuelle Restriktionen und Weltwissen bestimmt. Ein *onomasiologischer* Ansatz dagegen behandelt «Lage» in den genannten Ausdrücken als unterschiedliche Wörter, die die gleiche Gestalt haben. Eine dritte Lösung, im Sinne der Konstruktionsgrammatik, wäre es, die *ganzen* Ausdrücke als Lexeme mit unterschiedlicher Bedeutung zu betrachten (dieses Konzept wird im JAKOB-Lexikon mit dem Konzept der *Konstruktion* verfolgt).

Alle drei Ansätze sind im Grunde genommen unzureichend, weil im ersten Fall das Lexikoninventar sehr verarmt ist und weil die im Gebrauch existierenden Bedeutungen nicht erfasst werden können; im zweiten und dritten Fall geht die Tatsache verloren, dass die verschiedenen Bedeutungen von «Lage» doch zusammenhängen und aus einander hervorgegangen sind! Und die dahinter stehende Ausdruckseinheit (Lage) steht auch weiterhin für die Ausdrucksbildung zur Verfügung. Feilke bezeichnet es als eine zentrale Leistung der Sprache, mit einem kleinen Satz an Ausdruckselementen die enorme Komplexität möglicher Konzeptualisierungen fassen zu können (S. 130).

¹⁷Vgl. das *lexical priming* bei Hoey (2005), Kap. 6.12.

4.6 Verben, verbale Phraseologismen

Im folgenden Kapitel soll zuerst ein kurzer Überblick über einige Klassifikationssysteme für Verben vorgestellt werden. Verben haben in der Erzählanalyse eine wichtige Funktion; die Analyse der Erzählungen und die dafür verwendete Kodierung bauen in erster Linie auf den Verben als den Handlungen der Akteure auf. Verben haben in besonderem Masse die Eigenschaft, im Zusammenhang eines Satzes (Prädikation) Bedeutung herzustellen; insofern sind die Verbklassifikationen auch Typologien der Handlungen oder der Verbpattern. In diesem Sinne können auch einfache Verben als Phraseologismen angesehen werden, neben den eigentlichen verbalen Phraseologismen und Funktionsverbgefügen, die im zweiten Teil dieses Kapitels zur Sprache kommen.

4.6.1 Klassifikationssysteme für Verben

Die folgende Liste von Studien und Typologien zu Verben repräsentiert eine Auswahl mit unterschiedlichen theoretischen Hintergründen und Vorannahmen, aber auch mit verschiedenen Zielen. Das Interesse an diesen Systemen ist vor allem in der Frage begründet, welche Informationen für die Neugestaltung des JAKOB-Lexikons übernommen werden könnten und welche Typologien einen praktischen Nutzen für die Beschreibung von Lexikon-Items haben könnten. Die Liste erhebt keinen Anspruch auf Vollständigkeit.

- *English Verb Classes and Alternations*. Die Verbklassen von Levin (1993) stellen eine viel beachtete und oft aufgenommene, aber auch oft kritisierte semantische Klassifizierung der englischen Verben dar. Die fast 200 Verbklassen in 49 Kategorien basieren auf der Annahme, dass semantische und syntaktische Eigenschaften von Verben eng zusammenhängen. Als Beispiel gibt es eine Kategorie *Verbs of sending and carrying*, die die Verbklassen *Send, Slide, Bring and Take, Carry* und *Drive* enthält.
- *Handbuch deutscher Kommunikationsverben* (Harras, Winkler, Erb & Proost, 2004; Harras, Proost & Winkler, 2007). Eine semantische Beschreibung der *Sprechaktverben* des Deutschen, weiterer Verben des Sagens, der Gesprächsstrukturierung und Gesprächskennzeichnung, die in *Wortfeldern* (Paradigmen) eingeteilt und detailliert beschrieben werden.

- *Deutsche Verben* (Ballmer & Brennenstuhl, 1986): Eine sprachanalytische Untersuchung des deutschen Verbwortschatzes. Es handelt sich um ein Klassifikationssystem für Verben nach semantischen Kriterien. 8000 deutsche Verben sind in Klassen eingeteilt, die auf der nächsthöheren Ebene in Prozessmodellen organisiert werden.
- *Funktionsverbgefüge* (Pottelberge, 2001). Die (deutschen) Verbklassen *Verbonominale Konstruktionen* und *Funktionsverbgefüge* werden empirisch untersucht; eine klare Kategorisierung ist gemäss Autor nicht möglich, weshalb er die beiden Begriffe überhaupt in Frage stellt.
- Die semantische Klassifizierung der Verben im *OLIF-Lexikonmodell* stellt eine relativ einfach strukturierte und praktisch orientierte Merkmalsliste dar (OLIF-Eigenschaft *semType* für Verben), die im vorliegenden Projekt für die semantische Typisierung der Verben verwendet wird¹⁸ (McCormick, 2005).
- Schulte Im Walde (2005) präsentiert 43 semantische Verbklassen für das Deutsche (mit 168 Verben als Ankerbeispielen), die auf Intuition aufgebaut sind, und prüft dieses Klassifikation empirisch an Korpusdaten mit Hilfe eines Clusterverfahrens.
- Eine einfache Klassifikation der Verben und Funktionsverben nach semantischen Kriterien ist bei Helbig und Buscha (2000) zu finden.
- Das Kodiersystem der *Erzählanalyse JAKOB* gruppiert die Verben in 93 Klassen und 5 Dimensionen, auf dem Hintergrund des Bühnenmodells und psychoanalytischen und handlungstheoretischen Konzepten («Kategorienliste für handelnde Figuren und Handlungen», Boothe et al., 2002), siehe Kap. (7.1.1).

Daneben gibt es viele Studien zu *Verben* und *Funktionsverbgefügen*, die nicht eine Typologie zum Ziel haben, sondern die das Vorkommen und das Verhalten von verbalen Phänomenen in Korpora untersuchen. So beschreibt Imo (2007) die Konstruktionen, die von ausgewählten matrixsatzfähigen Verben im gesprochenen Deutsch gebildet werden, Overlach (2008) untersucht die Besonderheiten des Sprechens über Schmerzen (Schmerzkonstruktionen, allerdings nicht beschränkt auf Verben). Eine vergleichende Studie zu den Typologien von Ballmer und Brennenstuhl (1986), GermaNet (siehe

¹⁸siehe Kapitel (9.1.2).

Kap. 7.2.2), FrameNet¹⁹ und den semantischen Verbklassen von Levin präsentiert Schulte Im Walde (2005) und entwirft selber eine Typologie mit 43 semantischen Verbklassen (deutsche Verben). Das *Valenzwörterbuch deutscher Verben* (VALBU) listet die deutschen Verben alphabetisch und nach ihren Valenzmustern geordnet mit semantischen Informationen auf, kategorisiert die Verben aber nicht weiter in Gruppen (Schumacher, Kubczak, Schmidt & Ruiter, 2004).

Die umfangreiche Liste der *englischen Verbklassen* von Levin soll hier nicht detailliert aufgelistet werden (Levin, 1993). Stattdessen wird in Tabelle (4.1) die Verbklassifikation abgebildet, wie sie in GermaNet verwendet wird (Kunze & Lemnitzer, 2007). Diese lehnt sich an die Klassifikation von Levin an, besteht aber aus wesentlich weniger Klassen.²⁰

Tabelle 4.1: Verbklassen in GermaNet

Verbs of Possession	Besitzverben
Verbs of Location	Lokationsverben
Verbs of Emotion	Gefühlsverben
Social Verbs	Gesellschaftsverben
Verbs of Body	Körperfunktionsverben
Verbs of Cognition	Kognitionsverben
Verbs of Communication	Kommunikationsverben
Verbs of Competition	Konkurrenzverben
Verbs of Contact	Kontaktverben
Verbs of Creation	Schöpfungsverben
Verbs of Change	Veränderungsverben
Verbs of Consumption	Verbrauchsverben
Verbs of Perception	Perzeptionsverben
Stative Verbs	Zustandsverben

Die Verbklassen von Levin werden oft zitiert und als Grundlage für Klassifikationssysteme verwendet, werden aber auch oft kritisiert, z. B. von Schnorbusch (2005, S. 36), der bemängelt, dass kein fundiertes Inventar semantischer Eigenschaften im Hintergrund der Verbeinteilung steht und dass diese allein an die Intuition des Lesers appelliert. Schnorbusch stellt die Behauptung in Frage, dass semantische Klassen aus syntaktischem Ver-

¹⁹FrameNet in einer deutschen Umsetzung: SALSA II – The Saarbrücken Lexical Semantics Acquisition Project (Burchadt et al., 2009), siehe Kap. (7.2.3).

<http://www.coli.uni-saarland.de/projects/salsa/page.php?id=index> (15.11.2010).

²⁰http://www.sfs.uni-tuebingen.de/GermaNet/verb_classes.shtml (15.11.2010).

halten abgeleitet werden können. Hanks und Pustejovsky (2005) kritisieren an den Levin'schen Verbklassen ebenfalls die auf Intuition beruhende Typologisierung der Verben, die nicht empirisch fundiert und teilweise idiosynkratisch sei. Die Aussage von Levin, wonach das Auftreten eines Verbs zu einem grossen Teil durch seine Bedeutung bestimmt wird (*behavior of a verb to a large extent determined by its meaning*), wird von den beiden Autoren abgelehnt. Sie drehen die Argumentation um und sagen (auf Grund von Korpus-evidenz – Verbpatterns), dass die Bedeutung eines Verbs zu einem grossen Teil durch sein Auftreten bestimmt wird (*meaning of a verb to a large extent determined by its behavior*), weil das Verhalten des Verbs (Verbpattern) beobachtbar und verifizierbar ist (Hanks & Pustejovsky, 2005, S. 5). Ausserdem wird bemängelt, dass die Analysen der Verbbedeutung im Bereich des Satzes bleiben und nicht über den Satz hinausgehen.²¹

Das *Handbuch deutscher Kommunikationsverben* (HDK) teilt diese Verbggruppe in die folgenden *Paradigmen* oder *Wortfelder* ein: Allgemeine verba dicendi, Repräsentative, Direktive, Kommissive, Expressive, Deklarative etc. Die einzelnen Lexikoneinträge beschreiben detailliert Bedeutung, Argumentstruktur, spezielle Verwendungsweisen und mögliche synonymische Beziehungen in bestimmten Kontexten. Als Beispiel für den Aufbau der Typologie sei hier ein einzelnes *Wortfeld* zitiert, dessen Charakterisierung mit den folgenden Merkmalen vorgenommen wird (Harras et al., 2004, S. 122):²²

Das Wortfeld <Direktive.auffordern.fordern>

- Propositionaler Gehalt: Mitteilungsgehalt: P
- Geschehenstyp: Handlung
- Zeitbezug: zukünftig
- Rollenbezug: Hörer
- Einstellung von S zu P: S will: P
- Sprecherabsicht: S will: H tut: P
- Vorannahmen von S: nicht erwartbar: P

²¹Dieses Problem besteht auch bei der segmentweisen Kodierung in der JAKOB-Applikation und wird dort zu korrigieren versucht, indem thematische, situationale und linguistische Merkmale eines Segments auf das nächste übertragen und mit den neuen Merkmalen abgeglichen werden.

²²P = Prädikat, propositionaler Gehalt; S = Sprecher; H = Hörer.

- Modalität der Äusserung: nachdrücklich

Eine einzelne *Klasse* wird dann mit folgenden Merkmalen dargestellt, am Beispiel der Klasse *fordern*, die Verben wie «verlangen», «abverlangen», «wünschen», «einfordern» umfasst (gekürzt!):

- Syntaktische Umgebung: jemand fordert etwas (von jemandem) ...
- Bedeutung: jemand verlangt etwas mit Nachdruck von jemandem ...
- Verwendungsspezifk: Die Situationen, auf die mit *fordern* Bezug genommen wird, sind dadurch gekennzeichnet, dass ein Sprecher S einen Hörer H nachdrücklich auffordert, etwas Bestimmtes, P, zu tun ...
- Feldzugehörigkeit: <Dir.aufford.forder>, Bezugssituationstyp: Direktive, Untertyp: Aufforderungen, Untertyp: Forderungen.
- Synonyme: verlangen, abverlangen, wünschen, einfordern, (Forderungen stellen).
- Antonyme: keine.
- Kommentar: für alle Verben des Typs <Dir.aufford. forder> gemeinsame typbestimmende Bedeutungsaspekte.

Die vergleichsweise einfache semantische Klassifikation der Verben im OLIF-Lexikonformat²³ wird in der Tabelle (9.3) im Kapitel (9.1.2) beschrieben (vgl. McCormick, 2005). Sie umfasst 15 Kategorien, darunter emotion, event, verschiedene Arten von activity (unspecified, mental, physical), process, situation, perceptive etc. Die Tabelle (4.2) schliesslich zeigt ein einfaches Klassifikationsschema, wie es im *Leitfaden der deutschen Grammatik* (Helbig & Buscha, 2000, S. 35) verwendet wird.

Fazit zu den Klassifikationssystemen: Typologien über Typologien nützen wenig, weil sie meistens auf Intuition aufgebaut sind. Sie sagen mehr über die mentalen Konzepte aus, die in den Theorien abgebildet sind als über den realen Sprachgebrauch. Für das Lexikonprojekt zeigt es sich ausserdem, dass eine genaue und detaillierte Klassifikation der Verbeinträge zwar von theoretischem Interesse ist, aber für die Bedeutungsherstellung beim Verstehen (Textrezeption) nicht in gleichem hohem Grade nützlich ist.

²³Open Lexicon Interchange Format, URL: <http://www.olif.net/> (15.11.2010).

Tabelle 4.2: Verbklassifikation nach Helbig und Buscha (2000)

Tätigkeitsverben	Agens übt in aktiver Weise eine Tätigkeit aus.
Vorgangsverben	Subjekt (ohne Agens zu sein) erfährt eine Veränderung an sich, wird dabei verändert.
Zustandsverben	Zustand des Subjekts, der sich nicht ändert (statisch).
Aktionsarten	Verlaufsweise des Geschehens.
– Durative Verben	(imperfektiv) bezeichnen den Ablauf des Geschehens.
– Perfektive Verben	grenzen den Verlauf des Geschehens ein.
Funktionsverben	Liste der häufigsten FV, z. B. bringen, geben, erfahren.
Semantik der	Wichtigste Leistung: Ausdruck der Aktionsart:
Funktionsverben	Zustand (durativ), Zustandsveränderung (inchoativ), Bewirken von Zustand oder Veränderung (kausativ).

4.6.2 Funktionsverben und Funktionsverbgefüge

Funktionsverben werden in der Literatur unterschiedlich benannt, so gibt es die deutschen Bezeichnungen *Funktionsverben* (FV), *Nominalisierungsverben*, *verblasste Verben* und *Funktionsverbgefüge* (FVG, für den ganzen Ausdruck, nicht nur das Verb). Die Bezeichnungen meinen teilweise auch unterschiedliche Konzepte. Ein minimaler Konsens besteht darin, dass FVG Konstruktionen sind, die aus einem semantisch reduzierten Verb und einem Nomen bestehen, das eine Aktion oder einen Zustand bezeichnet (Langer, 2005). In englischer Sprache sind die Bezeichnungen *phrasal verbs*, *light verbs* oder *support verbs* gebräuchlich.

Das Vorkommen von Funktionsverben und die damit gebildeten Konstruktionen in grossen Korpora wird z. B. von Kamber (2008) untersucht, er analysiert das Vorkommen von als typisch bezeichneten Funktionsverbgefügen im Gegenwartssdeutsch. Im Rahmen eines DWDS-Projektes²⁴ analysieren Storrer (2007) und Hanks, Urbschat und Gehweiler (2006) Konstruktionen mit deutschen Funktionsverben. Hanks et al. (2006) machen einen Vorschlag für einen neuen Typ von Wörterbucheinträgen für (deutsche) *Funktionsverben*, begründet auf Korpusevidenzen (die gegenwärtigen monolingualen Wörterbücher für Deutsch sind grösstenteils Neuauflagen der Versionen der Vor-Korpus-Ära). Im wesentlichen bedeutet dieser neue Typ, dass Funktionsverben wie Idiome oder Konstruktionen behandelt werden. In der Studie werden Konstruktionen mit den Funktionsverben «leisten»,

²⁴Projekt «Kollokationen im Wörterbuch», auf Datengrundlage des DWDS (Digitales Wörterbuch der deutschen Sprache).

«erteilen», «hegen», «machen» und «bringen» untersucht. Wichtig ist, dass FVG nicht durch ein einfaches Verb als Synonym ersetzt werden können, wie es eine vereinfachende lexikographische Einteilung suggerieren könnte (vgl. Storrer, 2007, S. 164). Für das Funktionsverb «leisten» führen Hanks et al. (2006, S. 445) aufgrund der Suchergebnisse in den Korpora des DWDS vier unterschiedliche Kollokationsmuster auf:

1. «leisten» mit einem «event noun» als direktes Objekt macht 56% der Vorkommen aus. Es gibt aber dabei nur eine kleine Anzahl unterschiedlicher «event nouns», meistens Nomen, die eine soziale Interaktion zugunsten der Gesellschaft oder des Rezipienten (der nicht explizit erwähnt werden muss) bezeichnen, wie «Hilfe leisten». Oft sind diese Nomen auch Gegensatzpaare (z. B. Eid - Meineid, Gehorsam - Widerstand). Seltener vorkommende Nomen sind oft den prototypischen verwandt (alternations), «Unterstützung leisten» und «Assistenz leisten» können als Alternationen zu «Hilfe leisten» angesehen werden.
2. Reflexiver und transitiver Gebrauch des Verbes «sich etwas leisten» machen 31% der Vorkommen aus.
3. Das Verb hat als direktes Objekt ein abstraktes Nomen, das eine Quantität oder etwas Messbares bezeichnet: Diese Konstruktion macht 2% der Vorkommen aus. Bsp. «der Motor leistet 200 PS».
4. Das Verb hat als direktes Objekt ein indefinites Pronomen oder ein nominalisiertes Adjektiv. Dies trifft in 9% der Vorkommen zu (Bsp. «er leistet etwas», «er leistet nichts»). Gemäss den Autoren bezeichnet hier das indefinite Pronomen, wie oben in der ersten Bedeutung, einen Beitrag zur Gesellschaft.

Storrer (2007) betont auf Grund ihrer Studie, dass die Wahl zwischen FVG und einfachem Verb (z. B. zwischen Hilfe leisten und helfen) nicht so willkürlich ist wie oft in der Literatur angenommen wird. Für Zwecke der Textproduktion (Übersetzungssysteme oder «style checker») müssten deshalb systematisch die Entscheidungsfaktoren für die eine oder andere Variante auf Basis verfügbarer elektronischer Korpora erschlossen werden.

Kamber (2006, 2008) legt eine empirisch und korpuslinguistisch begründete Untersuchung von deutschen Funktionsverbgefügen vor. Er weist auf die oft nicht der Sprachrealität entsprechende Darstellung von FVG in Wörterbüchern hin:

Aufgrund dieses empirischen Defizits sind einige fundamentale Fragen bis heute nicht zufriedenstellend beantwortet worden, wie zum Beispiel: Welche sind die frequentesten FVG im alltäglichen Sprachgebrauch? Welche wichtigen FVG stehen (nicht) in einsprachigen Wörterbüchern des Deutschen? Welche wichtigen FVG stehen (nicht) in DaF-Grammatiken? usw. (Kamber, 2006, S. 109)

Er stellt deshalb *Monographien* von 10 Funktionsverbgefügen vor, mit dem Ziel einer adäquaten Beschreibung der wesentlichen Merkmale (Kamber, 2008): *setzen, stellen, stehen, nehmen, bringen, kommen, geraten, gehen, sich befinden, bleiben*.

Für Feilke dienen die Funktionsverben der *differenzierenden Ausdrucksbildung*; sie füllen eine lexikalische Lücke (Feilke, 1996, S. 142–147). Feilke teilt die Funktionsverbgefüge in die folgenden Gruppen ein (in drei konzentrischen Kreisen):

- Im Zentrum (im innersten Kreis) stehen die FVG im engeren Sinne, wie «zur Entscheidung kommen», «...bringen», «...stellen», «...stehen».
- Im zweiten Kreis befinden sich die *Nominalisierungsverbgefüge* (NVG). Beispiele: «jmdm Hilfe leisten», «jmdm Antwort geben». Ein Nominalabstraktum bildet hierbei die Kollokationsbasis.
- Im äussersten (dritten) Kreis schliesslich befinden sich die *Substantiv-Verb-Kollokationen* (SVK). Beispiele: «Geld abheben», «die Initiative ergreifen», «den Tisch decken». Die SVK sind lediglich charakterisiert durch häufige Kookurrenz der einzelnen Komponenten, nicht aber durch eine Motivation semantischer oder pragmatischer Art.

4.7 Formulaic language, Formelhaftes

Das folgende Kapitel schliesst an die Beschreibung der Phraseologismen an und überschneidet sich damit; der Fokus auf die Phraseologismen ist aber hier etwas anders gesetzt, nämlich auf die Ansicht, dass viele der Ausdrücke, die wir täglich verwenden, als Ganzes im mentalen Lexikon gespeichert sind und deshalb beim Sprechen und beim Verstehen nicht aus den Einzelkomponenten zusammengesetzt werden müssen. Die folgenden Abschnitte beleuchten also Phraseologismen speziell unter dem Aspekt der Formelhaftigkeit.

Ein grosser Teil unserer alltäglich verwendeten Sprache ist formelhaft (formulaic). Wray (2005) nennt als einführendes Beispiel den Ausdruck *Rice Krispies*: Bei einer Passantenbefragung (British TV, 1993) zeigten sich viele Leute erstaunt darüber, dass *Rice Krispies* aus Reis bestehen sollten. Entsprechend umschreibt Wray die Eigenschaften dieses und ähnlicher Ausdrücke so:

Words and word strings which appear to be processed without recourse to their lowest level of composition are termed *formulaic*. (S. 4)

Es gibt eine grosse Menge von Begriffen, die das Phänomen *formulaic language* bezeichnen; für die englische Sprache hat Wray (S. 9) eine Liste von Bezeichnungen aus der Literatur zusammengestellt (siehe Tabelle 4.3). Die Begriffe stellen jeweils andere Aspekte des Phänomens in den Vordergrund:

Tabelle 4.3: Begriffe für formelhafte Sprache (nach Wray, 2005)

Amalgams	Gambits	Preassembled speech
Automatic	Gestalt	Prefabricated routines and
Chunks	Holistic	patterns
Clichés	Holophrases	Ready-made expressions
Co-ordinate constructions	Idiomatic	Ready-made utterances
Collocations	Idioms	Rote
Composites	Irregular	Routine formulae
Conventionalized forms	Lexicalised phrases	Schemata
FEIs (fixed expressions including idioms)	Lexicalised sentence stems	Semi-preconstructed phrases that constitute single choices
Fixed expressions	Multiword units	Sentence builders
Formulaic language	Non-compositional	set-phrase
Formulaic speech	Non-computational	Stable and familiar expressions with specialized sub-senses
Formulas/formulae	Non-productive	Stereotypes
Fossilized forms	Petrification	Synthetic
Frozen phrases	Praxons	Unanalysed chunks of speech

Burger (2007) verwendet für die Eigenschaften von formelhafter Sprache den Begriff *pragmatische Festigkeit*, schränkt jedoch diese Bezeichnung weitgehend auf den Begriff der *Routineformeln* ein.

Vorgeformte Muster kommen nicht nur im Bereich einzelner Phraseologismen vor, sondern auch im grösseren Rahmen des Diskurses (vgl. Kapi-

tel 3.2.3). Günthner (2006b) zitiert die folgenden Autoren, die das Phänomen des Vorgeformten darstellen. Ono und Thompson (1995) beschreiben detaillierte, oberflächennahe und redundante *constructional schemata*; Hopper (1987) nennt diese Versatzstücke *pre-fabricated parts*, während sie bei Langacker *pre-packed units* heissen (Langacker, 1999). In der Konstruktionsgrammatik werden sie als *constructions* bezeichnet (Fillmore, Kay & O'Connor, 1988). Die Gattungsanalyse beschreibt ähnliche Phänomene im Kontext von Wissenssoziologie, Anthropologie und Linguistik. Die Gemeinsamkeit zwischen den genannten Konzepten besteht in der Ansicht, dass die sedimentierten Muster *verfestigte Schemata* sind, die auf Grund ihrer gestalthaften Qualität Projektionen über den weiteren Interaktionsverlauf erlauben (Günthner, 2006b, S. 2).

Die Grenze zwischen einem *formelhaften* und einem *nicht formelhaften* Ausdruck ist schwierig zu ziehen. Dazu kommt, dass nicht nur ganze Ausdrücke, sondern auch einzelne Wörter formelhaft verwendet werden können (Bsp. «verrückt» oder «normal»). Unterbrechungen im Redefluss fallen oft mit den Grenzen von Redeformeln zusammen, die in der interaktionalen Aushandlung präsentiert, verändert, getestet, zurückgezogen oder aktiviert werden (Hopper, 1987, S. 5).

Mündliche Kommunikation erfolgt unter Zeit- und Handlungsdruck. Auch aus dieser Sicht gibt es einleuchtende Argumente für die Verwendung von *formelhafter Sprache*. So schreiben Günthner und Bückler (2009):

Form und Funktion grammatischer Strukturen im mündlichen Gebrauch sind mit spezifischen interaktionalen und kognitiven Aspekten verbunden, die eng mit den Produktions- und Rezeptionsbedingungen der gesprochenen Sprache zusammenhängen. So ist eine [...] zentrale Beobachtung die, dass SprecherInnen – nicht zuletzt aufgrund der Tatsache, dass mündliche Kommunikation unter erheblichem Zeit- und Handlungsdruck abläuft – auf rekurrente, verfestigte grammatische Muster angewiesen sind, die Ressourcen für die Bewältigung kommunikativer Aufgaben darstellen. (S. 1)

Pawley und Syder (1983) weisen ebenfalls auf den grossen Anteil an vorgeformten Ausdrücken bei der Sprachproduktion hin und erwähnen gleichzeitig den Aspekt, dass diese Formeln redundant gespeichert sind:

speakers' mental lexicons do contain much more than just lexical primitives, namely hundreds of thousands of prefabricated items that *could* be productively assembled but are, as a result of

frequent encounter, redundantly stored and accessed. (Pawley & Syder, 1983, zit nach Gries, 2008, S. 18)

Formelhaftigkeit sieht Stein (2004) besonders in den Aufgaben der mündlichen Kommunikation: bei der sozialen Organisation der Interaktion, beim Kommunikationsmanagement und bei der eigentlichen Formulierungstätigkeit im Gespräch. Die Formulierungstätigkeit beinhaltet zwei grundlegende Arten von Aktivitäten: Kognitive Inhalte müssen erstens versprachlicht werden und zweitens muss der Sprecher Bezug nehmen auf bereits Versprachlichtes. Dazu gehören Verfahren und Indikatoren der Redebewertung und Redekommentierung, sowie Verfahren und Indikatoren der Bearbeitung, motiviert durch die rhetorischen Strategien ›Exemplifizierung‹, ›Generalisierung‹, ›Zusammenfassung‹, und ›Störungen‹ wie Rephrasierung, Wiederholung, Paraphrase, Distanzierung, Korrektur und Reformulierungen (Stein, 2004, S. 269).

Gebrauchsfertige sprachliche Einheiten sind besonders in *kritischen Formulierungsphasen* von Nutzen, um die Konstruktionskomponente zu entlasten. Die Simultaneität von Planung und Verbalisierung stellt oft zu hohe Ansprüche an die Sprecher und führt zu Überlastung (Stein, 2004, S. 277).

Bei Kuiper (2009) werden die formelhaften Strukturen *Phrasal Lexical Items* (PLI) genannt, und er entwickelt ein eigenes Regelsystem zur Funktion von PLI und ihren Komponenten. PLI sind Phrasen, Diskursteilnehmer kennen in der Regel mindestens zwei Wörter in der Phrase (lexikalisierte Konstituenten). Nicht jedes Wort muss in der PLI spezifiziert sein, die sog. *slots* einer PLI können frei belegbar oder mit *slot restrictions* versehen sein (syntaktische oder semantische Restriktionen). Daneben gibt es auch *bound words* (unikale Elemente); diese Wörter können nur in der entsprechenden PLI verwendet werden. In der PLI können ausserdem optionale Konstituenten vorgesehen sein, so dass die PLI ein flexibles und modifizierbares Gebilde ist, eine (kompositionale) Funktion ihrer Konstituenten (Kuiper, 2009, S. 19 f).

Formeln (formulae) im Sinne von Kuiper (2009, S. 21) sind PLI mit kontextabhängig eingeschränkten Gebrauchsbedingungen. Jeder Satz ist ein potentiell PLI, wenn er im Gebrauch verfestigt wird. In den PLI haben auch kreative Wortersetzungen oder -ergänzungen Platz. Kuiper spricht von *base form* und *variant*, die willkürlich als kreative oder künstlerische Varianten eingesetzt werden können, die aber auch aus Versprechern resultieren können. «The humourist exploits the resources of the phrasal lexicon.» (Kuiper, 2009, S. 191–198). Vgl. dazu auch die Theorie der *Norms and Exploitations* von (Hanks, in press), siehe Kapitel (6.11).

4.8 Fazit zum Kapitel Phraseologie

Für den Aufbau eines elektronischen (Konstruktions-)Lexikons zeigen sich in den Kapiteln zur Phraseologie wichtige Fragen und Probleme der Implementierung. Sollen alle möglichen Varianten (z. B. von «Lage», wie oben beschrieben) als eigenständige Einträge ins Lexikon aufgenommen werden, ergibt dies eine Flut von neuen Einträgen, die aber einen grossen Teil von Merkmalsausprägungen gemeinsam haben. Ein Vorteil ist, dass passende Lexikoneinträge schnell und direkt gefunden werden, der Nachteil hingegen ist der Lexikonumfang, der potentiell ins Unermessliche wächst, wenn das Lexikon über das Stadium der prototypischen Konstruktionen hinauskommt. Damit der oben (Beispiel 20, S. 97) angesprochene gemeinsame Hintergrund der Lexikoneinträge erhalten bleibt, nämlich die Gemeinsamkeiten der Ausdrücke mit «Lage» und die potentielle Kernbedeutung von «Lage», müssten alle diese Ausdrücke im Lexikon verknüpft werden (cross references, siehe Kap. 9.1.2), was wiederum – wie an der kleinen Liste der Lage-Ausdrücke ersichtlich – sofort zu einer Flut von expliziten Verknüpfungen führt.

Wird für den Lexikonaufbau nach dem oben erwähnten semasiologischen Konzept verfahren, ergibt sich das praktische Problem, dass der einzelne Lexikoneintrag mit optionalen Elementen und verschiedenen Bedeutungsaspekten überfrachtet ist, was eine adäquate Beschreibung der Konstruktionen verhindert. Möglicherweise ist es für den Lexikonaufbau wichtig zu entscheiden, ob das Lexikon für die Rezeption, für das Sprachverstehen, entwickelt wird, oder ob dabei die Produktion auch mitgedacht wird. Für das Verstehen kann mit einer kleineren Menge von Einträgen und Verknüpfungen ein praktisch annehmbares Resultat erzielt werden. Die Ableitung der Lexikoneinträge aus dem Korpusvorkommen zielt auch in Richtung eines Rezeptionslexikons. Für das konkret vorliegende Lexikonprojekt wird die erstgenannte Variante angestrebt, die kompatibel ist mit dem Begriff der Konstruktion als Einheit von Form und Bedeutung.

Phraseologie beinhaltet eine Menge von *Klassifizierungen* und *Typologien*, so z. B. für die Ordnung der Mehrwortausdrücke in Kategorien. Über die Abgrenzung der Kategorien besteht in der Literatur keine Einigkeit. Die phraseologischen Typologien sind meistens auf Intuition aufgebaut, was zwar von theoretischem Interesse ist, aber für die praktische Verwendung im Lexikon nicht in gleichem hohem Grade nützlich. Die Typologien sagen mehr über die mentalen Konzepte aus, die in den Theorien abgebildet sind als über den realen Sprachgebrauch.

Gemäss Burger können Phraseologismen verschiedenste *Funktionen* aus-

üben:

aber erstens ist es in einem konkreten Fall schwierig und nicht ohne interpretative Willkür möglich, einem Vorkommen eines Phraseologismus eine bestimmte Funktion zuzuordnen, und zweitens lassen sich die Funktionen nicht ohne Berücksichtigung der Textsorte bzw. der kommunikativen Situation beschreiben. (Burger, 2007, S. 156)

Die einleuchtendsten Funktionszuweisungen an Phraseologismen finden sich im Bereich der formelhaften Sprache; formelhafte Ausdrücke stehen oft im Dienst ganz bestimmter kommunikativer Aufgaben und Funktionen. Für das Lexikonprojekt ist diese Feststellung relevant; sie bedeutet, dass Phraseologismen in Bezug auf konkrete Texte und Gespräche, Text- und Gesprächsmuster und Textsorten untersucht werden müssen und nicht isoliert, d. h. das Lexikon muss die entsprechenden Kategorien und Merkmale bereitstellen, die eine Abbildung und Differenzierung verschiedener Funktionen ermöglichen.

Eine weitere wichtige Frage bezüglich der Charakterisierung von Phraseologismen wird von Seite der *Computerlinguistik* gestellt, betrifft allerdings weniger die Phraseologismen im Lexikon, sondern potentielle Phraseologismen in zu analysierenden Texten: Wie werden einzelne Phraseologismen oder zusammengehörige Ausdrücke im laufenden Text erkannt? Welche Wortketten bilden lexikalisch-semantische Einheiten und durch welche Art Verfahren kann dies festgestellt werden? Für das Lexikonprojekt kann die Frage präzisiert werden zur Frage, wie die im Lexikon erfassten MWE in den transkribierten Gesprächen erkannt werden können. Eine günstige Voraussetzung besteht darin, dass in der Erzählanalyse JAKOB die zu analysierenden Gesprächsausschnitte (Erzählungen) bei der elektronischen Erfassung manuell segmentiert werden (Subjekt-Prädikat-Einheiten, siehe Kapitel 2.3.2) und deshalb einen übersichtlichen Suchraum vorgeben. Das «Matchen» von Verbpatterns im Text mit Verbpatterns im Lexikon wird dadurch erleichtert.

Der im Kapitel zur Interaktionalen Linguistik vorgestellte Begriff der *Bedeutungskonstitution* bekommt unter phraseologischen Gesichtspunkten neue Akzente. Während in der Gesprächsanalyse die interaktiven und dialogischen Prozesse wichtig sind, die zur Herstellung von Bedeutung beitragen, wird der Fokus in der Phraseologie, Psycholinguistik und in der Kognitiven Linguistik auf die kognitiven Prozesse gelegt, wie dies am Beispiel der Metonymie gezeigt wird. Die kognitiven Aspekte der Bedeutungskonstitution ergänzen die interaktionalen und leiten über zum Thema der *dynamischen*

Bedeutungskonstitution, die nicht direkt im Lexikon abgebildet werden kann und deshalb über die Möglichkeiten des Lexikons hinausgeht. Begriffe wie Unterspezifizierung und Vagheit verweisen auf grundsätzliche Prinzipien der Kommunikation und ermöglichen zusammen mit den dialogischen Prinzipien ein Modell der Bedeutungskonstitution, wonach sich *unterspezifizierte Kernbedeutungen* im Lexikon befinden und bei der Sprachanwendung mit zusätzlichen lokalen Bedeutungskomponenten versehen werden, die situative Faktoren und Faktoren des Gesprächskontextes berücksichtigen.

5 Korpuslinguistik

Das folgende Kapitel stellt kurz die speziellen Methoden der Korpuslinguistik vor, die bereits im Kapitel zu Phraseologie und Computerlinguistik angesprochen wurden. Im Zusammenhang mit der Korpuslinguistik werden die Begriffe Kollokation, Kookkurrenz, Kontext und weitere erklärt, die in elektronisch aufbereiteten Korpora eine wichtige Bedeutung bekommen. Eine kurze Darstellung von statistischen Assoziationsmassen, die in der Korpuslinguistik und im vorliegenden Projekt verwendet werden, schliesst das Kapitel ab.

Die bisher vorgestellten Methoden zur Erforschung und Darstellung von natürlicher Sprache aus Gesprächen sind angewiesen auf den Zugang zu Ton- oder Videoaufnahmen kurzer Gesprächsausschnitte oder auf Transkripte dieser Gesprächsausschnitte. Text- und Gesprächsdaten sind in *Korpora* repräsentiert, die aus naheliegenden Gründen meistens auf schriftlichen Texten beruhen. Mittlerweile gibt es allerdings auch eine wachsende Zahl von Korpora transkribierter *Gespräche*, die teilweise öffentlich zugänglich sind. Die Theorien und Methoden der *Korpuslinguistik* können dazu beitragen, die in kleinen Gesprächsausschnitten mit gesprächsanalytischen Methoden gefundenen Erkenntnisse auf grössere Textmengen zu beziehen und Besonderheiten des Sprachgebrauchs auch empirisch zu belegen. Der englische Korpuslinguist John Sinclair leistete Pionierarbeit für eine funktionale Sicht auf sprachliche Phänomene, indem er auf Grund von Korpusuntersuchungen die *Musterhaftigkeit* von Sprache postulierte. Lemnitzer und Zinsmeister (2006) charakterisieren die Korpuslinguistik wie folgt:

Als Korpuslinguistik bezeichnet man die Beschreibung von Äusserungen natürlicher Sprachen, ihrer Elemente und Strukturen, und die darauf aufbauende Theoriebildung auf der Grundlage von Analysen authentischer Texte, die in Korpora zusammengefasst sind. (S. 7)

Die Korpuslinguistik hat viele Annahmen der klassischen Linguistik in Frage gestellt, weil mit den Korpusdaten der Sprachgebrauch empirisch überprüft werden kann und weil die Ergebnisse aus Korpusanalysen zeigen, dass der Sprachgebrauch oft nicht den intuitiven Sprecherurteilen der

Linguisten entspricht. Korpora stehen in einem gewissen Widerspruch zur Grammatik. Dies zeigt sich darin, dass Texte (z. B. aus der Chat- und Email-Kommunikation) und vor allem Gesprächstexte oft nicht wohlgeformt sind und «Fehler» enthalten (und trotzdem vom Leser oder Hörer problemlos verstanden werden). Einerseits gibt es also in den Korpora Sätze, die der Grammatik widersprechen, und andererseits gibt es viele grammatikalische Konstruktionen, die nicht im Korpus gefunden werden können (Lemnitzer & Zinsmeister, 2006, S. 24). Dies demonstriert die bekannte Diskrepanz zwischen *Kompetenz* und *Performanz*.

Der *Kontextualismus* postuliert, dass linguistische Erkenntnisse aus dem Sprachgebrauch erschlossen werden sollten. Wichtige Vertreter des Kontextualismus sind die Korpuslinguisten John Rupert Firth und John Sinclair. Sprachliche Äußerungen und deren linguistische Aspekte sind Funktionen des *sprachlichen und nicht-sprachlichen Kontextes*, in dem diese Äußerungen stehen (Lemnitzer & Zinsmeister, 2006, S. 28). Im Rahmen des Kontextualismus gelten die folgenden, für die Korpuslinguistik wichtigen Definitionen der Begriffe Kontext und Kotext (Lemnitzer & Zinsmeister, 2006, S. 29–30):

Definition Kontext: Der Kontext einer Äußerung ist die Summe der unmittelbaren Rahmenbedingungen einer Sprachhandlung (kultureller und situativer Kontext). Der Begriff geht in der Linguistik auf John Rupert Firth zurück.

Definition Kotext: Der Kotext einer linguistischen Einheit ist die Menge der linguistischen Einheiten, die im gleichen Text verwendet werden (Kotext von Wörtern und Sätzen z. B. auf den Ebenen der Phonetik, Morphologie, Syntax).

Definition Kollokation 1: Im Paradigma des Kontextualismus wird unter Kollokation das wiederholte faktische Miteinander-Vorkommen von zwei oder mehr Wörtern oder lexikalischen Einheiten verstanden.

Im Kontextualismus wird also jedes gemeinsame Vorkommen von zwei Wörtern in gleichen Kotext als Kollokation bezeichnet. Dieser Kollokationsbegriff ist aber zu weit gefasst und wenig nützlich (siehe auch Kap. 4.2.2). Lemnitzer und Zinsmeister (2006, S. 147) führen deshalb zusätzlich den Begriff der *Kookkurrenz* ein und definieren den Kollokationsbegriff neu. Kollokation soll nicht nur statistische Kriterien des gemeinsamen Auftretens berücksichtigen, sondern die Kollokation soll eine interne Bedeutungsstruktur haben. Auch Granger und Paquot (2008, S. 42) plädieren dafür, dass der

Begriff *Kollokation* nicht für statistische Kookkurrenzen verwendet wird, sondern für die traditionelle Bedeutung als gebrauchsbasierte lexikalisch eingeschränkte Wortkombination. Die entsprechend angepassten Definitionen lauten demgemäss (Lemnitzer & Zinsmeister, 2006, S. 147):

Definition Kookkurrenz: Gemeinsames Vorkommen zweier Wörter in einem gemeinsamen Kontext, der als Textfenster einer bestimmten Länge festgelegt wird (vgl. Steyer, 2004a).

Definition Kollokation 2: Die Kollokation weist neben den Kriterien aus Definition 1) (faktisches Miteinander-Vorkommen von zwei oder mehr Wörtern) zusätzlich eine interne Bedeutungsstruktur auf, in Form einer Hierarchie zwischen Kollokationsbasis und Kollokator und die Komponenten einer Kollokation stehen in einer syntaktischen Beziehung zueinander.

Definition Kolligation: Paare sprachlicher Einheiten, deren Zusammenhang durch die Bezeichnung ihrer syntaktischen Kategorien und deren Relationen gekennzeichnet ist (z. B. Funktionsverbgefüge).

5.1 Statistische Assoziationsmasse in der Korpuslinguistik

Kookkurrenz – das gemeinsame Vorkommen von Wörtern – kann in elektronischen Korpora mittels statistischer Assoziationsmasse ermittelt werden. Diese Masse setzen die Kookkurrenzhäufigkeiten und die Häufigkeiten der einzelnen Wörter einander gegenüber und berechnen so eine Bewertungszahl, deren Höhe etwas über die Stärke der Assoziation, d. h. über die Häufigkeit des gemeinsamen Auftretens der beteiligten Wörter aussagt. Seit den 1990er Jahren wurden verschiedene Assoziationsmasse entwickelt (Heid, 2007, S. 1042); eine Übersicht und einen Vergleich von über 30 dieser Verfahren präsentiert Evert (2005). Im Korpusanalysenwerkzeug *sketch engine* (SKE) werden verschiedene Assoziationsmasse angeboten, mit denen die Stärke des Zusammenhangs von Kollokationskandidaten beziffert werden kann (Beispiele im Durchführungsteil, Kapitel 9.2).¹ Die Zusammenhangsmasse können einerseits verwendet werden, um *text mining* zu betreiben, zum Beispiel um Hypothesen über das Kollokationsverhalten von Ausdrücken in Korpora nachzuprüfen, wie dies mit Hilfe der *sketch*

¹Eine ausführliche Beschreibung der *sketch engine* erfolgt im Kapitel (8.3.2).

engine im vorliegenden Projekt unternommen wird. Im Bereich der Computerlinguistik werden aber die statistischen Masse auch verwendet, um *Mehrwortausdrücke* in grossen Korpora automatisch zu identifizieren und zu klassifizieren (Heid, 2008).

Im Folgenden werden zwei Assoziationsmasse erwähnt, die in der *sketch engine* zur Anwendung kommen und dadurch im vorliegenden Projekt eine praktische Bedeutung haben. Es handelt sich einerseits um *Mutual Information* (MI) und andererseits um den *logDice-Score*.

5.1.1 Mutual Information

Mutual Information (MI) als statistisches Mass für die Kollokationsstärke wurde 1990 von Church und Hanks eingeführt (Church & Hanks, 1990). MI kann direkt aus maschinenlesbaren Korpora (mit linguistischen Informationen versehen) berechnet werden und wird wie folgt dargestellt:

$$MI(x, y) = \log_2 \frac{P(x, y)}{P(x)P(y)}$$

MI vergleicht die Wahrscheinlichkeit des gemeinsamen Auftretens der Wörter x und y (gemeinsame Wahrscheinlichkeit) mit den Wahrscheinlichkeiten des unabhängigen Auftretens der Wörter x und y (Zufallswahrscheinlichkeit). Wenn also eine tatsächliche Beziehung zwischen x und y besteht, wird die gemeinsame Wahrscheinlichkeit $P(x, y)$ wesentlich grösser sein als die Zufallswahrscheinlichkeit $P(x) P(y)$, und damit wird $I(x, y) > 0$. Wenn keine erkennbare Beziehung zwischen x und y besteht, ist $I(x, y)$ um Null herum, und wenn x und y in komplementärer Distribution vorkommen, d. h. wenn das eine Wort im Kontext das andere ausschliesst, ergibt sich $I(x, y) < 0$ (Church & Hanks, 1990, S. 23).

5.1.2 logDice Score

Kritikpunkte am Assoziationsmass MI betreffen den Umstand, dass der Wert von *Mutual Information* keine stabile Grösse ist, sondern vom jeweiligen Korpus abhängt und deshalb nicht für den Vergleich von Kollokationsparametern zwischen verschiedenen Korpora geeignet ist. Rychlý (2008) präsentiert ein Assoziationsmass *logDice*, das besser geeignet scheint, Assoziationsmasse zwischen verschiedenen Korpora zu vergleichen.

$$\logDice = 14 + \log_2 D = 14 + \log_2 \frac{2f_{xy}}{f_x + f_y}$$

Das Assoziationsmass *logDice* hat die folgenden Eigenschaften (Rychlý, 2008):

- Das theoretische Maximum des Scores ist 14 Punkte (wenn alle Vorkommen von x mit allen Vorkommen von y zusammenfallen). Übliche Werte für gute Kollokationskandidaten betragen weniger als 10 Punkte.
- Negative Werte zeigen an, dass es keine statistisch signifikante Kookkurrenz von x und y gibt.
- 1 Punkt Unterschied im Score entspricht ungefähr einer Verdoppelung des gemeinsamen Vorkommens von x und y.
- Das logDice-Mass hängt nicht von der Gesamtgrösse eines Korpus ab, weil relative Häufigkeiten von xy im Verhältnis zu x und y verglichen werden.

In der *sketch engine* stehen für die Auswertung von Kollokationskandidaten beide Assoziationsmasse zur Verfügung. Beim Vergleich der beiden Masse in der korpuslinguistischen Arbeit hat das logDice-Mass auf den ersten Anblick die plausibleren Ergebnisse geliefert (dieser Befund wurde allerdings nicht systematisch überprüft).

5.2 Fazit zum Kapitel Korpuslinguistik

Im Kapitel (4.4) wurde bereits die Tatsache angesprochen, dass die einfachen Wortassoziationsmasse nur *Wortpaare* berücksichtigen; nützlich wäre aber auch das Erkennen von grösseren Einheiten, wie dies unter der Bezeichnung *multiple Kookkurrenzcluster* im Korpusanalysesystem COSMAS² möglich ist, wie Duffner und Näf (2006) im folgenden Beispiel beschreiben:

Beispiel: Zu <Herz> erscheint nicht einfach nur das Kollokat <schlagen>, sondern darüber hinaus speziellere multiple Cluster wie <Herz+schlagen+höher> oder <Herz+ schlagen+ aufgehört>, usw. Dabei wird für jede Kombination die Frequenz und

²URL: <http://www.ids-mannheim.de/cosmas2/> (15.12.2010)

für die Hauptkookkurrenz <schlagen> der LLR-Wert (logarithmic likelihood ratio), ein Mass für die «Exklusivität» der Beziehung zwischen zwei Argumenten, angegeben. (Duffner & Näf, 2006, S. 14)

Für das vorliegende Projekt werden die einfachen Masse *logDice* und *Mutual Information* verwendet, da diese in der sketch engine angeboten werden, und weil die hauseigenen Korpora für die vorausgesetzten linguistischen und statistischen Analyseprozeduren in die sketch engine integriert werden können. Die Assoziationsmasse zeigen höchst interessante Phänomene auf und sind für Korpusanalysen unverzichtbar, weil intuitive Schlüsse aus reinen Häufigkeitsangaben in den Konkordanzlisten ungenaue und irreführende Ergebnisse liefern.

6 Konstruktionsgrammatiken

Nachdem im Kapitel zur Narrativik und Erzählanalyse eine Sicht auf Wörter und Ausdrücke gezeigt wurde, die die Interpretation von bestimmten Sprachproduktionen zum Ziel hat, wurde im Kapitel zur Interaktionalen Linguistik auf die interaktionellen und kommunikativen Funktionen von Gesprächspraktiken und Äusserungsformen fokussiert, die im Sprachgebrauch an bestimmten Ausdrucksformen und Abläufen festgemacht werden. Im Kapitel zur Phraseologie erfolgte dann eine mehr formale Sicht auf einen Aspekt dieser Phänomene, nämlich die Gruppierung von Wörtern zu mehr oder weniger komplexen Mehrwortausdrücken mit speziellen Eigenschaften und Funktionen. Im folgenden Kapitel zur *Konstruktionsgrammatik* stehen nun wieder Wörter und ihre Verbindungen zu Ausdrücken im Zentrum des Interesses; diesmal ist der Fokus aber auf den Begriff der *Konstruktion* als einer Einheit von Form und Bedeutung gelegt, die im mentalen Lexikon gespeichert ist und aus einem Einzelwort oder aus Wortverbindungen beliebiger Komplexität bestehen kann.

Bevor die Konstruktionsgrammatik ausführlicher beschrieben wird, werden *funktionale Grammatiktheorien*, zu denen auch die Konstruktionsgrammatik gehört, in einem Überblick kurz vorgestellt. Dann wird der Begriff der *Konstruktion* an einigen Beispielen exemplifiziert, und die Hauptrichtungen der Konstruktionsgrammatik und ihre Grundannahmen werden erläutert. Neuere Entwicklungen umfassen sowohl Bemühungen, die ursprünglich syntaxlastige Konstruktionsgrammatik mit gesprächslinguistischen und interaktional-linguistischen Theorien in Verbindung zu bringen als auch Erweiterungen der konstruktionsgrammatischen Theorien mit semantischen und pragmatischen Aspekten und Erweiterungen des Konstruktionsbegriffs auf die Satz- und Diskursebene. Das Phänomen der Grammatikalisierung wird im Zusammenhang mit Konstruktionen angesprochen, und schliesslich werden einige Formalismen zur Darstellung von Konstruktionen vorgestellt. Damit wird auch die Verbindung gemacht zum vorliegenden Lexikonprojekt, dessen zentrales Thema die Konstruktion als Idee, die Formalisierung der Konstruktionen und ihre Abbildung im Lexikon ist.

Unter dem Oberbegriff der Konstruktionsgrammatik stelle ich zwei weitere Sprachtheorien vor, die eine funktionale Sicht auf die Sprachverwendung

vertreten und grosse Ähnlichkeiten mit Konstruktionsgrammatiken aufweisen, nämlich die *Corpus Pattern Analysis*, die Gebrauchsmuster von Verben untersucht und die Theorie des *Lexical Priming*, die die individuelle Sprachentwicklung als Funktion des mehr oder weniger frequenten Gebrauchs von Wörtern und Ausdrücken in Kontexten postuliert.

6.1 Funktionale Grammatiktheorien

Wilhelm von Humboldt betrachtete die Sprache als Organ, als eine Art Körperteil des Menschen, während Karl Bühler in seinem Organon-Modell die Sprache als Instrument bezeichnete. Auch Ludwig Wittgenstein sprach von der Sprache als einem Instrument; von ihm stammt die *Gebrauchstheorie der Bedeutung*:

Die Bedeutung eines Wortes bzw. eines sprachlichen Zeichens wird als die Regel seines Gebrauchs in der Sprache L definiert. Das bedeutet, dass der Sprecher, der die Regel des Gebrauchs eines Wortes in der Sprache L kennt, somit die Bedeutung des Wortes beherrscht. (Smirnova & Mortelmans, 2010, S. 5)

Für Ferdinand de Saussure jedoch war die Sprache ein *System*, womit er den *Strukturalismus* begründete und die Unterscheidung von *langue* (Sprachsystem) und *parole* (Sprechfähigkeit) einführte (Smirnova & Mortelmans, 2010, S. 8).

Die *Dependenztheorien* gehen auf Lucien Tesnière zurück; sie messen dem *Verb* als regierendem Element des Satzes grosse Bedeutung bei und bilden den Hintergrund zu verschiedenen verb- und satzbasierten Grammatiken und Sprachtheorien (vgl. die Verbpattern der CPA, Kap. 6.11). Das Verb ist das Zentrum des Satzes oder der Szene, die Mitspieler sind Aktanten und Zirkumstanten (in heutigen Dependenzmodellen *Ergänzungen* und *Angaben* genannt. Tesnière beschreibt die Aktanten eines Verbes als «Akteure in einem kleinen Drama».

1. prime actant, fait l'action.
2. second actant, supporte l'action.
3. tiers actant, l'action se fait à son profit ou à son detriment. (Tesnière 1965, nach Götz-Votteler, 2007, S. 37)

Die *Bühnenmetapher* mit der für den Satz zentralen Bedeutung des Verbs wird auch in der Erzählanalyse JAKOB aufgenommen; auch hier sind die

Verben das Zentrum und beschreiben das Handeln der Akteure in der Erzählung (die Erzählungen werden in Segmente gegliedert, die zeigen, wie die Akteure handeln: «Wer tut was wie?», vgl. Kap. 2.3.2):

Valenzmodelle stellen, ähnlich wie die *Dependenzmodelle*, das Verb in den Mittelpunkt. Unterschiede zwischen *Dependenz* und *Valenz* bestehen darin, dass sich *Valenz* nur auf die Relationen zwischen dem Verb und seinen Ergänzungen bezieht, während *Dependenz* alle Abhängigkeitsbeziehungen innerhalb eines Satzes beschreibt. In den gängigen *Valenzmodellen* wird zwischen verschiedenen Ebenen der *Valenz* unterschieden; so postuliert Helbig (1992) drei *Valenzebenen*: logische, semantische und syntaktische *Valenz*. Für das vorliegende Projekt ist vor allem die semantische und die syntaktische *Valenz* relevant sowie die Verbindung zwischen den beiden. Für die Beschreibung *semantischer Valenz* erwähnt Götz-Votteler (2007, S. 38) exemplarisch vier Möglichkeiten:

1. *Semantische Rollen* (semantic roles), auch Tiefenkasus, thematische oder Theta-Rollen genannt, gehen auf die von Charles Fillmore entwickelte *Kasusgrammatik* zurück. Zu einem Verb gehören thematische Rollen wie z. B. *Agens*, *Patiens*, *Instrument*. Anzahl und Art dieser Rollen differieren stark in den verschiedenen Studien zu diesem Konzept.
2. *Semantische Komponenten* (semantic components) spezifizieren die semantischen Eigenschaften der Nomen, die als Komplemente des Verbs auftreten.¹ Beispiele: *animate*, *human*, *abstract*, *document* etc. Die Beispiele zeigen, dass es verschiedene Abstraktionsebenen dieser Eigenschaften gibt. In der *Corpus Pattern Analysis* werden die *semantischen Typen* der Komplemente in einer umfangreichen Ontologie abgebildet (Kap. 6.11).
3. *Semantische Kategorien*: Als weitere Möglichkeit, die eng mit der vorher erwähnten zusammenhängt, wird die Klassifikation der syntaktischen Komplemente in semantische Kategorien aufgeführt, wie sie im VALBU eingeführt wird. Die Kategorien entsprechen in etwa den semantischen Typen.
4. Ebenfalls im VALBU gibt es zusätzlich eine verbspezifische Beschreibung der beteiligten «Teilnehmer» (verb-specific description of par-

¹Diese semantischen Eigenschaften werden in der Terminologie der generativen Grammatik auch als *Selektionsrestriktionen* bezeichnet.

ticipants). Auch dies ist eine Ergänzung zum semantischen Typ und beschreibt weitere Eigenschaften auf ein bestimmtes Verb bezogen.

Semantische Valenzmodelle sind eng verbunden mit der Framesemantik. Je genauer die Eigenschaften möglicher Komplemente eines Verbs beschrieben werden, desto mehr kommen extralinguistische Faktoren wie Diskurskontext, Situation, Wissen um Kommunikationspraktiken und Weltwissen ins Spiel und schränken damit die theoretisch unbeschränkte Kombinationsmöglichkeit von Wörtern ein, wie sie in der *Generativen Grammatik* (GG) angenommen wird: «Sprache wird dabei als die potentiell infinite Menge von wohlgeformten grammatischen Sätzen verstanden, die nach fest definierten Regeln aus einer finiten Anzahl von sprachlichen Elementen generiert werden.» (Smirnova & Mortelmans, 2010, S. 12).

Funktionale Grammatiktheorien entwickelten sich als Gegenströmung zur Generativen Grammatik und betonen die Rolle der zwischenmenschlichen Kommunikation, die Sprache ist Werkzeug und Kommunikationsmittel. D. h. die Umstände der menschlichen Sprechttätigkeit, wie Situation und Kontext sind wichtige Parameter. Funktionale Theorien betrachten sowohl die kommunikativen als auch die kognitiven Funktionen als wichtig für eine Grammatik. Zu den funktionalen Grammatiken gehören z. B. die *Funktionale Grammatik*, die vom Linguisten Simon Dik in den 1970er Jahren begründet wurde, oder die *systemisch-funktionale Grammatik/Linguistik*, in den 1960er Jahren von Michael A. K. Halliday begründet, die bis heute Beachtung findet (Halliday & Matthiessen, 2004).²

Die *Grammatikalisierung* kann als eine funktionale Grammatiktheorie verstanden werden. Viele grammatikalische Items sind ursprünglich aus lexikalischen Items entstanden (vgl. Kap. 6.8). In einem ähnlichen Sinn postuliert die *Emergent Grammar* (Hopper, 1987), dass Grammatikregeln und syntaktische Strukturen *emergent* beim Sprachgebrauch entstehen. Hopper grenzt diese Art Grammatik von der *a priori Grammar* ab (vgl. Kap. 3.3.5).

²Systemic-Functional Linguistics (SFL) is a theory of language centred around the notion of language function. While SFL accounts for the syntactic structure of language, it places the function of language as central (what language does, and how it does it), in preference to more structural approaches, which place the elements of language and their combinations as central. SFL starts at social context, and looks at how language both acts upon, and is constrained by, this social context.

A central notion is «stratification», such that language is analysed in terms of four strata: Context, Semantics, Lexico-Grammar and Phonology-Graphology.

Context concerns the Field (what is going on), Tenor (the social roles and relationships between the participants), and the Mode (aspects of the channel of communication, e.g., monologic/dialogic, spoken/written, +/- visual-contact, etc.).

(URL: <http://www.isfla.org/Systemics/definition.html> (15.11.2010))

Einen informativen und nützlichen Überblick über funktionale Grammatiktheorien geben Smirnova und Mortelmans (2010). Auch die *Kognitive Grammatik* (Langacker, 2008) und die *Konstruktionsgrammatik*, gehören zu den funktionalen Grammatiken.

6.2 Konstruktionsgrammatik

Bei der *Konstruktionsgrammatik* (*Construction Grammar* (CxG)) handelt es sich weniger um ein einheitliches Theoriegebäude als vielmehr um eine ganze «Familie» von Theorien. Zu den ausgearbeiteten, mehr oder weniger formalisierten Konstruktionsgrammatiken³ zählen unter anderem die folgenden:

- Die in Berkeley entstandene Richtung der Konstruktionsgrammatik um Charles Fillmore und Paul Kay (Fillmore et al., 1988), eine eher formalistisch orientierte Variante.
- Die kognitiv-linguistisch orientierte Ausprägung um Adele Goldberg (1995) und George Lakoff (1987).
- die Radical Construction Grammar von William A. Croft (2001).
- die Kognitive Grammatik von Ronald Langacker (2008).
- Weitere kognitiv-linguistische Zugänge, wie der Denkansatz der *Mental Spaces* (Fauconnier, 1994).
- Interaktional ausgerichtete Ausprägungen der Konstruktionsgrammatik, vor allem im deutschen Sprachraum (z. B. Auer, 2006; Deppermann, 2006a; Günthner & Bucker, 2009; Hopper, 1987; Imo, 2007).

Diese Einteilung hat allerdings mehr historisch-orientierenden Charakter; aktuelle CxG-Projekte lassen sich nicht so einfach zuordnen. Die Theorien der CxG sind in den 1980er Jahren im Umfeld der kognitiven Linguistik entstanden. Vorläufer der CxG sind Fillmores *case grammar* (Fillmore, 1968) und Lakoffs Begriff der «Gestalt» (Lakoff, 1987). In den letzten 20 Jahren entstand viel Literatur zum Phänomen der Konstruktion; dass *Konstruktionen* einen wichtigen Stellenwert in der Sprache einnehmen, ist mittlerweile

³Der Begriff *Grammatik* darf hier nicht zu eng gesehen werden. Konstruktionsgrammatiken sind linguistische Theorien, die nicht nur grammatische Phänomene umfassen, sondern ebenso semantische und pragmatische.

breit anerkannt. Die Umsetzung dieser Theorien in Forschungsprojekten und praktischen Anwendungen ist allerdings noch weniger verbreitet, und besonders im deutschen Sprachraum gibt es erst seit jüngerer Zeit Einführungen zum Thema Konstruktionsgrammatik (Deppermann, 2006a; Feilke, 2007; Fischer & Stefanowitsch, 2007b; Günthner & Imo, 2006a; Stefanowitsch & Fischer, 2008).

Die wichtigste Annahme, die allen Ausprägungen der Konstruktionsgrammatik gemeinsam ist, ist die, dass sprachliches Wissen und *Bedeutung* auf der Ebene von Konstruktionen repräsentiert sind und vollzogen werden, d. h. aus Ausdrücken verschiedenster Komplexität, nicht nur aus Einzelwörtern. Konstruktionen sind *Form-Bedeutungspaare* und bestehen aus Ausdrücken oder Phraseologismen, die ganz oder teilweise lexikalisiert sind. Die Konstruktionen stehen in unmittelbarem Bezug zu ihren Funktionen im Sprachgebrauch, sie sind *usage-based* (Croft, 2001; Goldberg, 1995).⁴ In den linguistischen Theorien generativer Prägung besteht das Sprachwissen demgegenüber aus einer Sammlung von abstrakten Regeln, die universale Gültigkeit haben und die bei der Sprachproduktion angewendet und mit Einträgen aus dem Lexikon «gefüllt» werden. Die Konstruktionsgrammatik wendet sich damit gegen das abstrakte Regelsystem und die Prämisse der *Wohlgeformtheit* von Sätzen der generativen Grammatik und untersucht die Sprache unter tatsächlichen Gebrauchsbedingungen, d. h. wie sie von Muttersprachlern gelernt wurde und gesprochen wird.

6.2.1 Was sind Konstruktionen

Die folgende Erzählung zeigt beispielhaft, dass Konstruktionen wesentliche Anteile unseres alltäglichen Sprechens ausmachen.

(21) Transkript Wilma: «Typisch Frau»⁵

01 W: ich kann mich schLECHT (-)so (-) dann äussern;
 02 I: sie haben diesem freund NICht gesagt- eh du,
 03 das ist (--) das ist dann also MEIne Sache,
 04 ob ich ein Kind will oder nicht oder

⁴ «Das Usage-based model (gebrauchsbasiertes Modell) geht von der Vorstellung aus, dass sprachliches Wissen durch einen induktiven, bottom-up gerichteten Lernprozess erworben wird. Neben redundanter Speicherung grammatischer Information sind demnach auch Generalisierungen zugelassen, die sich aus Verallgemeinerungen des Lerners über sich wiederholende Ereignisse der Sprachverwendung (usage events) ergeben» (Quelle: Wikipedia, <http://de.wikipedia.org/wiki/Konstruktionsgrammatik>).

⁵ Diese Gesprächsausschnitt ist im Unterschied zu allen andern nach dem GAT-Standard transkribiert (Selting et al., 1998).

```

05 W:  ich habe es ihm dann SPäter eeeh::: noch gesagt;
06 I:  ja
07 W:  eh::- er habe mich dann SCHAUrig verlEtzt mit dieser bemerkung;
08      und dann hat er dann (--) aber AUCh wieder sehr wie mein
        vater auch=eh:- gefunden
09      was ICH jetzt für ein ZEUGs mache- (---)
10      und das sei TYpisch FRAU so ZeugS immer so auseinander
11      zu beineln und zu an(h)alys(h)ieren (h)oder(h) (-)
12      und SCHON haben wir natürlich wieder krAch gehabt
13      miteinander[ oder?
14 I:  [ja=ja=ja

```

Die in dieser Erzählung verwendeten Phraseologismen zeigen anschaulich, welches Format Konstruktionen haben können und wie sie verwendet werden. Konstruktionen können mindestens in den folgenden Ausdrücken mühelos erkannt werden: «Zeug», «ein Zeug machen», «Zeug analysieren und auseinander beineln», «typisch Frau», «Krach haben miteinander». Aber der Begriff der Konstruktion umfasst auch alltäglichere Phänomene und Wortkombinationen, und im Rahmen der verschiedenen Ausprägungen von Konstruktionsgrammatiken werden unter dem Begriff *Konstruktion* ganz unterschiedliche Gebilde verstanden. Die folgende tabellarische Zusammenstellung (Tab. 6.1) zeigt einige Beispiele von «Konstruktion» genannten Ausdrücken, die auf verschiedenen Ebenen der Abstraktion und der Granularität angesiedelt sind (die Aufstellung soll Beispiele zeigen und erhebt keinen Anspruch auf Vollständigkeit).

Tabelle 6.1: Konstruktionen

Darstellung	Bezeichnung/Beispiel
[V Subj OBJ OBJ ₂]	Ditransitiv-Konstruktion (Goldberg, 1995)
CAUSE-RECEIVE<agt rec pat>	do.
[WH VP]	Interrogativ-Konstruktion (Stefanowitsch, 2010)
[was X VERB {betrifft, angeht}]	<i>was-X-betrifft</i> -Konstruktion (Birkner, 2008)
Incredulity-Response-Construction	Bsp. «Du und studieren?!» (Feilke, 2007)
haben Angst um [jmdn, etwas]	<Angst haben um>-Konstruktion

6.3 Hauptrichtungen der Konstruktionsgrammatik

Die Arbeiten von Fillmore et al. (1988) beschreiben eine Variante der Konstruktionsgrammatik, die sich generativen Grammatikformalismen annähert und in der *Framesemantik* von Charles Fillmore begründet ist (Fillmore, 2003). Die Grundsätze dieser Richtung lassen sich wie folgt zusammenfassen (vgl. Östman & Fried, 2005a, S. 1): Die CxG ist generativ, d. h. sie soll linguistische Konstruktionen nicht nur im Sprachgebrauch beschreiben und analysieren, sondern sie soll ebenso auch die Generierung beliebiger neuer Äusserungen mit diesen Regeln ermöglichen. Die verschiedenen linguistischen Ebenen der Phonologie, Morphologie, Syntax, Semantik und Pragmatik sollen in die Theorie und in die Formalismen integriert sein, und die Grammatik soll eine Universalgrammatik sein, d. h. sie soll universale Aspekte der Sprache (aller Sprachen) abbilden. Die Konstruktionen sollten ausserdem konsistent sein mit unserem Wissen über Kognition und soziale Interaktion.

Die verwendeten Formalismen sind teilweise den in der *Head Driven Phrase Structure Grammar* (HPSG) verwendeten Merkmalstrukturen ähnlich; so verbindet z. B. Müller in seinen Arbeiten Konstruktionsgrammatik und HPSG (Müller, 2006, 2007, 2008).

Die *kognitiv-linguistische Ausprägung* der Konstruktionsgrammatik verbindet sich vor allem mit den Namen Adele Goldberg und John Lakoff (Goldberg, 1995; Lakoff, 1987). Goldberg legt Wert darauf, dass Aussagen über die Sprachrealität nur auf Grundlage empirischen Materials aus dem Sprachgebrauch gemacht werden können.⁶

Thematische Frames sind kognitive Schemata, die die Teilnehmer (Akteure, semantische Rollen) eines Ereignisses/einer Situation beschreiben. Diese Rollen werden in Form von Argumenten im *Prädikatframe* (Satzmuster) instantiiert; die Argumente beziehen sich auf den Frame oder das Satzmuster des entsprechenden Lexikoneintrags. Prädikatframes beziehen sich also auf

⁶«Die Kognitive Linguistik wendet sich gegen eine modulare Auffassung der Sprachfähigkeit, d. h. gegen die Hypothese, dass die Sprache – insbesondere die Syntax – ein autonomes System bilde, das mittels eines spezialisierten angeborenen Moduls erworben werde und nach eigenen Regeln und Prinzipien funktioniere. Stattdessen nimmt die Kognitive Linguistik an, dass die Sprachfähigkeit auf allgemeine kognitive und perzeptuelle Fähigkeiten zurückzuführen ist und zumindest teilweise durch sie motiviert ist» (Quelle: Homepage der Deutschen Gesellschaft für Kognitive Linguistik, DKGL),

URL: <http://www.dgkl-gcla.de/dgkl.html> (20.11.2010) .

die Oberflächenstruktur einer bestimmten Sprache, während sich thematische Frames eher mit übereinzelsprachlichen Strukturen beschäftigen.⁷ Der Unterschied zwischen thematischen und Prädikatsframes geht auf Goldberg (1995) zurück, die *Argumentrollen* und *Teilnehmerrollen* (argument roles and participant roles) unterscheidet, die sich auf die Konstruktion oder auf den Frame (eines bestimmten Verbausdrucks) beziehen.

Die *Radical Construction Grammar* von William Croft (2001) legt ebenfalls Wert darauf, dass das Modell usage-based ist, betrachtet Konstruktionen darüber hinaus überhaupt als einzige Ebene syntaktischen Wissens und hat enge Bezüge zur *Cognitive Grammar* von Ronald Langacker (Langacker, 1987, 2009), welche ebenfalls zur Familie der Konstruktionsgrammatiken gerechnet werden kann. Konstruktionen verhalten sich wie lexikalische Items und können als lexikalische Items gedacht werden. Dabei gibt es ein *Kontinuum* von syntaktischen zu lexikalisierten Konstruktionen, diese sind mehr oder weniger schematisch, mehr oder weniger lexikalisiert (im Beispiel (22) die schematische Transitivkonstruktion (a) und die lexikalisierte Konstruktion (b)).

- (22) a. [NP_{NOM} V_{FIN} NP_{AKK}]
 b. [[REFLDET]ins Bockshorn jagen lassen]

Für Croft sind *Konstruktionen* (Abb. 6.1) die einzigen primitiven Elemente der Sprachstruktur; alle anderen Kategorien werden aus Konstruktionen hergeleitet (Croft, 2001, S. 18). Crofts Postulat bedeutet, dass ein Konstruktionslexikon aus vielen Redundanzen besteht, was gemäss Croft der psycholinguistischen Erkenntnis entspricht, dass die mentale Repräsentation von Konstruktionen von ihrer Gebrauchshäufigkeit abhängig ist (S. 28). Dass die Grundlage für das Lernen und die Anwendung von Sprachwissen nicht in allgemeinen und abstrakten Regeln besteht, wurde z. B. von Tomasello (2007) gezeigt. Das Erlernen einer Sprache erfolgt über das Lernen von Konstruktionen. Das Kind lernt zuerst einzelne konkrete Ausdrücke wie «Katze» oder «Hund», geht dann über zu sogenannten *Holophrasen* (englisches Beispiel bei Tomasello: «Lemme-see») und *Pivot-Konstruktionen*⁸ und verwendet und kombiniert schliesslich die formelhaften Phrasen. Diese

⁷ «Thematic frames are cognitive schemata specifying the type of participants involved in the situation described by an event. These participants can be instantiated in the form of arguments in the predicate frames assigned to the lexical units linked to that event. Therefore, predicate frames are lexical constructs belonging to a particular language, but they are constructed from the interlingual thematic frames located in the ontology» (Pascual et al., 2008, S. 12).

⁸ Apokoinu oder Drehsatz genannt: «Das hat so weh getan, hat das.»

werden angewendet, bevor das analytische Regelwissen vorhanden ist. Dies wurde sowohl für den Erstspracherwerb, als auch für den Zweitspracherwerb nachgewiesen.

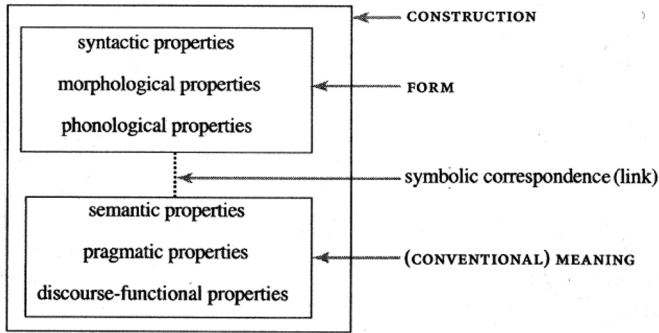


Abbildung 6.1: Symbolische Struktur einer Konstruktion (Croft, 2001)

Siepmann (2006) bringt die Ideen von Tomasello mit den Sprachtheorien des britischen Kontextualismus in Verbindung, wie sie z. B. von Hoey mit der Theorie des *Lexical Priming* vertreten werden (Hoey, 2005), wonach Sprecher ihre «lokale» Grammatik ad hoc und in ständiger dynamischer Bewegung entwickeln (Kap. 6.12). In der *Radical Construction Grammar* sind diese «lokalen Lexikoneinträge» (Konstruktionen) die kleinsten Einheiten der Sprachanwendung, während Kategorien wie *Nomen* oder *Adjektiv* lediglich Abstraktionen der nachträglichen Sprachbetrachtung sind (Imo, 2007, S. 27).

Der *Begriff der Konstruktion* ist sehr weit gefasst und in den verschiedenen Modellen der Konstruktionsgrammatik unterschiedlich definiert (Fischer & Stefanowitsch, 2007a). Konstruktionen enthalten grundsätzlich Informationen aus allen linguistischen Beschreibungsebenen: Syntax, Morphologie, Prosodie, Semantik, Pragmatik, Diskurs (Sequenzialität). Viele gesprächsanalytische Untersuchungen arbeiten implizit oder explizit mit dem Begriff der Konstruktion, wie die Bezeichnungen «Vorgeformtes», «formelhafte Wendungen», «feste Wortverbindungen» und ähnliche zeigen (Gülich, 2007b; Stein, 2004). Diese Begriffe stellen eine Verbindung zur *Phraseologie* her, die ebenfalls zwischen Grammatik und Lexikon angesiedelt ist, und die deshalb viele Bezüge zur CxG hat (Kap. 4).

Die *Kognitive Grammatik* von Ronald Langacker (2008) hat viele Gemeinsamkeiten mit den bereits erwähnten Konstruktionsgrammatiken und

ist die Grammatiktheorie, die am meisten in der kognitiven Linguistik verankert ist; die basalen menschlichen Erfahrungen des Sich-Bewegens, der sinnlichen Erfahrung und des Handelns in der Welt sind die Grundlagen für die sprachliche Darstellung und Rezeption. Die Kognitive Grammatik hat drei Hauptanliegen (Langacker, 2008):

1. «Grammar is meaningful in its own right»: Die symbolischen Strukturen (Konstruktionen) sind hinreichend für die Charakterisierung der komplexen Konstruktionen und Patterns. Die symbolischen Strukturen sind Paare von Form und Bedeutung, schematische Konstrukte, die ein semantisches und ein phonologisches Element besitzen sowie eine symbolische Beziehung zwischen den beiden Elementen.
2. Die Grammatik ist ein integraler *Bestandteil der menschlichen Kognition*, und nicht ein Spezialmodul. Die generellen kognitiven Fähigkeiten umfassen Abstraktion, Schematisierung, Kategorisierung und die Verwaltung der Aufmerksamkeitsressourcen.
3. Grammatik bildet *fundamentale Aspekte der menschlichen Erfahrung* ab. Konzeptuelle archetypische Rollen, begründet in der menschlichen Aktion und Interaktion in und mit der Welt werden auf Erfahrungshintergrund postuliert: agent, patient, instrument, experiencer, mover etc. sind prototypische Elemente für Konstruktionen.

Gemäss Langacker sind lexikalische Einheiten mehr oder weniger verfestigte Konstruktionen (prefabricated units), (Langacker, 1998, zit. nach Weber, 2010, S. 153). Das Lexikon ist a priori ein individuelles (mentales) Lexikon eines einzelnen Sprechers, das in einem zweidimensionalen Raum zwischen Schematizität und symbolischer Komplexität aufgebaut ist. Durch *entrenchment* (mentale Verfestigung durch häufige Wiederholung) werden komplexe Ausdrücke als Ganzheiten gespeichert, während die eher schematischen Konstruktionen durch variable Elemente «gefüllt» werden können. Weber zitiert als Beispiel die Konstruktion «Schwalbe» und Variationen: «Eine Schwalbe macht noch keinen Sommer» ist abgebildet im verfestigten Schema (6.3). Das Konzept Schwalbe ist eine typische Komponente des Konzepts Sommer (Weber, 2010, S. 159).

- (23) *Eine Schwalbe macht noch keinen Sommer.*
 Ein_{nom} typischer_Teil_{nom} macht noch kein_{akk} Ganzes_{akk}

6.4 Grundannahmen Konstruktionsgrammatik

Deppermann (2006a, S. 48 f.) fasst die Grundannahmen der Konstruktionsgrammatik mit *Fokus auf Gesprächsanalyse und interaktionale Linguistik* wie folgt zusammen:

Konstruktionen als umfassender Beschreibungsrahmen sprachlichen Wissens

Konstruktionen sind Kollokationen, feste Wendungen, Idiome und Metaphern, die lexikalisch ganz oder teilweise spezifiziert sind. Unregelmässige Konstruktionen stehen im Zentrum; die CxG ist eine Bottom Up-Grammatik. Äusserungen werden von Sprechern nicht regelgeleitet aus atomaren Einheiten aufgebaut, sondern als syntaktische Ganzheiten gelernt. Die Bedeutung von lexikalischen Einheiten existiert nicht isoliert, sondern entsteht erst im kollokationalen Kontext von ganzen Ausdrücken oder Konstruktionen.

Symbolic Thesis

Konstruktionen besitzen eine eigene Semantik, ihre Bedeutung ist nicht modular aufgebaut, sondern Konstruktionen sind *Einheiten von Form und Bedeutung*, «pairings of form and meaning» (Croft & Cruse, 2004). Konstruktionen sind nicht nur semantisch, sondern auch pragmatisch an Konstruktionskontexte gebunden. Es gibt keine Trennung von Syntax und Lexikon, sondern einen kontinuierlichen Übergang zwischen beiden Bereichen. Dies ergibt sich aus der Annahme, dass sowohl Wörter als auch komplexe grammatische Konstruktionen aus Form und Bedeutung bestehen.

Empirischer Ansatz: Das Usage-Based Model

Der empirische Zugang ist vor allem aus Sicht der interaktionalen Linguistik und Gesprächsanalyse wichtig und macht die Konstruktionsgrammatik reizvoll; es wird mit realen Gesprächsdaten gearbeitet, der Einzelfall wird genau dokumentiert. Paradoxerweise gibt es aber in der CxG wenig empirische Untersuchungen mit mündlicher Textdaten.

Weitere Merkmale der CxG

Günthner und Imo (2006b, S. 5) stellen in Übereinstimmung und Ergänzung zu den erwähnten drei Punkten die folgenden Merkmale der Konstruktions-

grammatik ins Zentrum:

1. Die Struktur von Sprache kann mit *Form-Bedeutungs-Paaren* (Konstruktionen) beschrieben werden, diese bilden ein systematisch verwobenes strukturiertes Inventar.
2. Sprache ist nicht ein autonomes, kognitives Symbolsystem, sondern ist eingebettet in die allgemeinen kognitiven Fähigkeiten (Langacker, 2008).
3. Die Grammatik ist *nicht-modular und nicht-derivational*.
4. Grammatik ist nicht angeboren, sondern *erlernt* auf dem Hintergrund von kognitiven Mechanismen und interaktionalen, auch sprachenspezifischen, Anforderungen.
5. Alle *Ebenen der Sprachbeschreibung* sind gleichwertig für die Beschreibung von Konstruktionen (Syntax, Semantik, Pragmatik).

Fragen, die bei allen Autoren immer wieder auftauchen, betreffen den *Status der Konstruktionen*: Entsprechen Konstruktionen der psychologischen Realität oder sind sie theoretische Konstrukte? Wie werden Konstruktionen identifiziert? Kontrovers diskutiert wird auch immer wieder die Frage, wie weit Konstruktionen verfestigt und lexikalisiert sind und wie weit sie «emergent» im Zuge des Diskurses entstehen.

6.5 Konstruktionsgrammatik und Interaktionale Linguistik

Fischer (2007) äussert sich zum Verhältnis von CxG und Interaktionaler Linguistik und zeigt, dass die Annahmen und Postulate der Konstruktionsgrammatik kompatibel sind mit den Kategorien der Gesprächsanalyse, Konversationsanalyse und Interaktionalen Linguistik. Die Aufnahme pragmatischer Information in die Eigenschaften einer Konstruktion legt diese Verbindung nahe; speziell einleuchtend ist dies z. B. für idiomatische Ausdrücke, die für die Festlegung der Bedeutung stark auf kontextuelle pragmatische Informationen angewiesen sind, wie z. B. «die Hand geben». Die Gesprächsanalyse kann also Hinweise darauf geben und zeigen, wie und wo Konstruktionen gebraucht werden. Fischer (2007, S. 135) spricht in diesem Zusammenhang von einer aus Grammatik und Interaktion bestehenden *Interaktionalen Konstruktionsgrammatik*, wie sie auch von weiteren Autoren

postuliert wird (Aijmer & Stenström, 2005; Fried & Östman, 2005; Imo, 2007). Gesprächsanalyse kann aufzeigen, unter welchen Kontextbedingungen Konstruktionen verwendet werden und was ihre interaktiven Funktionen sind.⁹ Pragmatische Gebrauchs- und Kontextbedingungen schränken die Verwendungsmöglichkeiten im tatsächlichen Sprachgebrauch oft mehr ein als man es intuitiv erwarten würde (Fischer, 2007, S. 138). Fischer plädiert für eine Nutzung der Konversationsanalyse, «um Teilnehmerinnenkategorien zu identifizieren, die die Verwendung von bestimmten Konstruktionen bedingen» (S. 141). Die Theorien der CxG und ihre Formalisierungen sind aber im allgemeinen eher syntaxlastig; lexikosemantische und pragmatische Aspekte sind noch weniger erforscht (Fischer, 2007, S. 140).

Es gibt nicht nur deutschsprachige Autoren, die dafür plädieren, Ideen und Methoden der Konstruktionsgrammatik und der Interaktionalen Linguistik (Konversationsanalyse) kombiniert zu verwenden; Fried und Östman (2005, S. 1754) betonen die gemeinsamen Interessen an der *Interaktion*. Gemeinsamkeiten sind z. B. darin zu finden, dass mit beiden Methoden Sprachfragmente (*chunks of language*) untersucht werden, ausserdem haben Konstruktionen auf der Diskursebene ähnliche Funktionen wie der *turn-type*. Konstruktionen und konversationelle Muster sind in Netzwerken von gegenseitigen Bezügen organisiert, und beide Zugänge sind sowohl *non-reduktionistisch* als auch *maximalistisch*, d. h. beide legen Wert darauf, dass ein volles Verständnis der sprachlichen Bedeutung und der linguistischen Strukturen nicht nur auf Form, Bedeutung oder Funktion allein reduziert werden kann (Fried & Östman, 2005, S. 1755). Unterschiede bestehen in der *Bedeutung des Kontextes* in Bezug auf die Äusserungen oder Konstruktionen. In der Konversationsanalyse ist der Kontext eher interaktional bestimmt, d. h. durch die die Gesprächspartner und die Situation; in der CxG ist Kontext eher ein linguistischer Begriff. Kontext wird hier eher in Frames oder Schemata verortet (siehe auch Kap. 3.3).

Zum Verhältnis der Konstruktionsgrammatik zur Interaktionalen Linguistik merkt Auer an, dass in der IAL die Emergenz syntaktischer Strukturen in Echtzeit untersucht wird, während die CxG sich auf die Rekonstruktion dieser Strukturen bezieht; die Idee des lokalen Managements, der Aushandlung emergenter Strukturen steht damit den mehr oder weniger dekontextualisierten kognitiven Konstruktionen gegenüber (Auer, 2006, S. 292). Aber auch die Gesprächsanalyse hat kognitive Komponenten, dazu gehören z. B. Wissen zu sequentiellen Ablaufschemata und kommunikatives

⁹z. B. für die Identifikation von *Frames*, die als semantischer Bezugsrahmen für Konstruktionen verwendet werden.

Gattungswissen. Das heisst, dass in den Konstruktionsmerkmalen Wissen über die Verwendung der Konstruktion im interaktionalen Kontext enthalten sein muss! Pragmatische Attribute gehören deshalb ebenso zu einer Konstruktion wie die semantischen (wie Genre, Register, Topic, siehe Kap. 7.3)¹⁰ (Auer, 2006, S. 294).

Eine Untersuchung von Deppermann und Elstermann (2008) zu *Konstruktionen mit dem Verb «verstehen»* zeigt wichtige und ergiebige mögliche Vorgehensweisen für Untersuchungen zu Konstruktionen auf dem Hintergrund der Interaktionalen Linguistik und Konversationsanalyse. Eine Kritik erwächst in diesem Zusammenhang gegenüber *Definitionen* von Konstruktionen, die in vielen Fällen kontextfrei beschrieben werden. Gemäss Deppermann und Elstermann (2008, S. 107) ist ein solches Verständnis des Begriffs Konstruktion nicht haltbar und wird auch empirisch, auf Grund von Korpusdaten, häufig widerlegt.

Auch die *Gesprochene-Sprache-Forschung* (im deutschen Sprachraum) greift die Ideen der Konstruktionsgrammatik zunehmend auf und integriert sie in die Theoriebildung zu einer *Grammatik der gesprochenen Sprache* (u. a. Auer, 2006; Deppermann, 2006a; Fischer & Stefanowitsch, 2007b; Günthner & Imo, 2006b; Günthner, 2007a; Imo, 2007; Stefanowitsch & Fischer, 2008).

Es gibt verschiedene Ansätze, Konversationsanalyse und Konstruktionsgrammatik (CxG) auf theoretischer Ebene zu verbinden. Dies erfordert allerdings von Konversationsanalytikern das Zugeständnis, dass die Bedeutung einer Äusserung nicht nur in der Interaktion konstituiert wird, sondern dass Bedeutung auch in *Konstruktionen* und mehr oder weniger fixierten Phrasen (d. h. in einem gewissen Sinne kodiert) vorliegen kann (Aijmer & Stenström, 2005). Studien zum Thema «formelhafte Sprache» (formulaic language) zeigen, dass tatsächlich ein grosser Teil des Ausdruckrepertoires in Gesprächen aus vorgeformten oder partiell vorgeformten sprachlichen Versatzstücken besteht (Dausendschön-Gay & Gülich, 2007; Stein, 2004; Wray, 2005). Fried und Östman (2005) zeigen in ihrem Artikel zu «Konstruktionsgrammatik und Interaktionaler Linguistik», wie Konstruktionen (linguistische Strukturen) im Gespräch von den Teilnehmenden («Mitgliedern») gemeinsam kreiert werden können (co-construction), dass aber das Resultat, die gemeinsam erarbeitete Struktur, meistens einem kodifizierten abstrakten Muster oder einer Verkettung von mehreren solchen Mustern entspricht:

This allows us to talk about the concrete coconstructing of un-

¹⁰Vgl. auch Deppermanns Studien zum *deontischen Infinitiv* (Deppermann, 2006b, 2007)

derstanding in discourse, while at the same time accepting that the result of co-constructing may correspond to codified abstract patterns (or, possibly, a concatenation of several such patterns). (Fried & Östman, 2005, S. 1775)

Die Konversationsanalyse zeigt den Kontext auf, in dem Konstruktionen verwendet werden können, CA dagegen zeigt die interaktiven Funktionen der Konstruktion auf. Fischer (2007, S. 141) plädiert (undogmatisch) für eine Vereinbarkeit von CA und externen Wissensressourcen (wie Konstruktionswissen).

6.6 Neuere Entwicklungen der Konstruktionsgrammatik

Neben den «etablierten» Hauptrichtungen der CxG gibt es verschiedene weiter führende Entwicklungen, so z. B. die «embodied construction grammar» (Bergen & Chang, 2005) oder die «fluid construction grammar» (Trijp, 2008), wie sie im Sammelband von Östman und Fried (2005b) beschrieben werden.

In einem Artikel mit dem Titel «From usage to grammar: the mind's response to repetition» zeigt Bybee (2006), dass die speziellen Gebrauchsbedingungen der Sprache die kognitive Repräsentation beeinflussen. Sie unterscheidet für ihre Darstellung «idioms», «prefabs» und «collocations». Als «prefabs» (prefabricated word combinations) bezeichnet sie konventionalisierte Wortsequenzen, die im Gegensatz zu Idiomen semantisch transparent sind.¹¹ Dazu zählen feste Wendungen, aber auch Verbgefüge und Verb-Präposition-Verbindungen. Prefabs machen gemäss Bybee einen Anteil von ca. 55% des gesprochenen und des geschriebenen Diskurses aus (Bybee, 2006, S. 713). Häufige Verwendung von Redewendungen und Wörtern stärkt deren Repräsentation im Gedächtnis, so dass sie sich Änderungen eher widersetzen und als feste Wendungen erhalten bleiben (Bybee, 2006, S. 715).

Eine gebrauchsbasierte Grammatiktheorie muss sich auf *allgemeine kognitive Prinzipien* beziehen: Auf die Wichtigkeit der *Wiederholung* beim Einschleifen (entrenchment) von neuromotorischen Mustern, das Verwenden von *Ähnlichkeit* in der Kategorisierung, und die *Generalisierung* über ähnlichen Mustern. Diese Prozesse, kombiniert mit den Funktionen der Sprache im sozialen Kontext, wie Referenzherstellung, Kohärenzherstellung

¹¹im Jakob-Lexikon wird dafür die Eigenschaft «set-phrase» verwendet.

und Organisation des Turn-taking, erklären Grammatik als die (ritualisierte) Verfestigung von sprachlichen Routinen (Bybee, 2006, S. 730). Auch Belica (2008) erklärt die Entstehung regelhafter Sprachphänomene aus dem Sprachgebrauch. Er betont die *emergente Natur* lexikalisch-semantischer Relationen und betrachtet das Phänomen Sprache aus einer emergentistischen Perspektive: Alles Regelhafte und Konventionelle in der Sprache ist ein Epiphänomen des Sprachgebrauchs, der von den Sprachteilnehmern laufend ausgehandelt wird (S. 1–2). *Ähnlichkeit* ist ein wichtiges Gegenkonzept zur *Identität*: Situationen und Kontexte, in denen wir uns zurechtfinden müssen, sind nie vollkommen identisch. Deshalb ist die phylo- und ontogenetisch gegebene Fähigkeit sehr wichtig, auch vage und verschwommene Ähnlichkeiten wahrzunehmen (Belica, 2008, S. 3). Dies gilt nicht nur für die Wahrnehmung, sondern auch für das Meinen und Verstehen in der Kommunikation (vgl. die «semantische Plastizität» bei Feilke, 1996, S. 62).

Die *Korpuslinguistik* (siehe Kap.5) hat den Ideen der Konstruktionsgrammatik weiteren Auftrieb gegeben. Die Bedeutung von Kollokationspatterns ergibt sich aus der Analyse ihres Vorkommens und ihrer praktischen Verwendung in grossen Textkorpora (Hanks, 2004). Besonders *John Sinclair* hat durch seine Arbeiten der Bedeutung von sprachlicher *Musterhaftigkeit* durch die Konkordanzanalyse in grossen Korpora neue Perspektiven erschlossen (Feilke, 2007; Hanks, 2008b).

Für die Konzeption des Lexikons übernehmen wir einstweilen die folgenden wichtigen Annahmen der Konstruktionsgrammatik: Konstruktionen sind Form-Bedeutungspaare in einem Syntax-Lexikon-Kontinuum, d. h. ein bestimmter Ausdruck ist immer mit einer bestimmten Bedeutung gekoppelt, die Ausdrücke können aus einem oder mehreren Wörtern bestehen, der Ausdruck kann mehr oder weniger variabel und damit ganz oder teilweise lexikalisiert sein. Mit den realen Gesprächsdaten wird auf die gesprochene Sprache fokussiert, was der Untersuchung der vorliegenden Transkripte und Erzählungen entgegenkommt. Die Syntax der Texte ist nicht wohlgeformt und enthält Ausdrücke aus dem Schweizerdeutschen.

6.7 Von der Konstruktionsgrammatik zur Diskursgrammatik

Östman (2005) schlägt vor, die Konstruktionsgrammatik über den Bereich des Satzes oder der Äusserung hinaus zu erweitern auf die Ebene des Diskurses. Dieser Ansatz ermöglicht eine präzisere Beschreibung der Konstruk-

tionen und widerspricht nicht den konstruktionsgrammatischen Prinzipien, weil ja die pragmatische Komponente zur Definition der Konstruktion gehört. Östman sieht CxG und Framesemantik als Basis einer vollwertigen Sprachtheorie an, die aus Fillmore's Case Grammar entstanden ist (Fillmore, 1968). Wenn Diskurseinheiten als Konstruktionen betrachtet werden, gibt es weitere, über die Semantik hinausgehende Aspekte der Pragmatik und der Diskursanalyse zu beachten (Östman, 2005, S. 123). Auch Imo (2007) schlägt vor, den Begriff der *Konstruktion* vom Wort über den Satz hinaus bis zum *Diskurs* auszudehnen, weil jede Konstruktion semantische, funktionale und diskursspezifische Eigenschaften besitzt. Östman schlägt für Diskurskonstruktionen (Diskursmuster) die Eigenschaften *Texttyp* und *Genre* vor, die Form, Bedeutung und Funktion eines Textes/Diskurses beschreiben (vgl. Kap.9.1.3):

- *Genres* sind kontextuelle Vorgaben einer *Gesprächssorte*, wie sie z. B. in Rezepten, Todesanzeigen, Tischgesprächen, Märchen, Arztgesprächen zum Ausdruck kommen.
- *Texttypen* dagegen charakterisieren die Eigenschaften einer *Diskurseinheit*, die z. B. argumentativ, narrativ, instruktiv, expositorisch oder deskriptiv sein kann.

Die vorgeschlagenen Eigenschaften gelten sowohl für geschriebenen Text als auch für den gesprochenen Diskurs (Östman, 2005, S. 130). In den Kapiteln (7.3 und 9.1.3) werden diese Ideen weiter ausgeführt und für die praktische Verwendung angepasst.

Auch Linells (2009) Auffassung der Konstruktionsgrammatik kann als Hinwendung zu einer Diskursgrammatik verstanden werden. Er erweitert die Konstruktionsgrammatik um den Begriff der *Dialogizität* und entwickelt ein eigenes Modell für die Repräsentation von Konstruktionen, das *SPRAC-Modell* (S. 137). Er fokussiert dabei auf die Interaktion im Diskurs, genauer auf den lokalen sequenziellen Kontext. Äusserungen haben immer responsive und projektive Funktionen (vgl. Günthner & Bücker, 2009). Konstruktionen haben ihren funktionalen Ursprung in *Gesprächspraktiken*, die auf sequenziellen und funktionalen, zum Teil auch auf formalen Kriterien beruhen. Als Beispiel zitiert Linell die Gesprächspraktik der *formulation*. Formulations¹² beziehen sich auf das vorhergehende Gesprächsthema und sind eine Art zusammenfassendes Fazit des Sprecher A über die vorange-

¹²Übersetzungsmöglichkeiten bewegen sich zwischen «Formulierung», «Reformulierung» und «Interpretation».

hende Äusserung des Sprechers B.¹³ Eine Formulierung leitet gemäss Linell eine dreiteilige Sequenzstruktur (Konstruktion) ein; sie ist erstens (*other-responsive*, weil sie die vorhergehende Äusserung von Sprecher B zusammenfasst und die Implikationen beschreibt, mit der Möglichkeit von A, eigene strategische Relevanzen zu setzen. Sie ist zweitens *other-projective*, weil mit der Kommentierung von B's Äusserungen mögliche weitere Kommentare, Bestätigung oder Ablehnung erwartet werden können, und sie sind drittens *self-projective*, weil sie für A einen dritten Turn projizieren, abhängig von B's Reaktion im zweiten Turn (Linell, 2009, S. 305).

Konstruktionen haben zusätzlich zur linguistischen Form und Semantik eine *externe Syntax* mit drei Aspekten:

1. Welche Bedingungen an die vorausgehenden Sequenzen zeichnen eine Konstruktion aus (eine Art Erkennungsmerkmal für die Konstruktion)?
2. Welche Bedingungen eröffnet die Konstruktion für die nachfolgenden Sequenzen?
3. Was sind häufige oder bevorzugte gleichzeitig auftretende Ressourcen? (Linell, 2009, S. 312)

Konstruktionen sind *Methoden*, «um etwas in der Interaktion zu tun» oder sprachlich zu handeln («to do something in interaction», S. 313). Die Konstruktion als Methode stellt einerseits eine formale Vorgabe für die Äusserungsgestaltung zur Verfügung (als linguistische Struktur), und andererseits trägt die Konstruktion etwas zur Bedeutungsherstellung bei, ist ein gestalterisches Merkmal für die Interpretation, ein Startpunkt für Projektionen etc. Zu einer Konstruktion kann es z. B. gehören, dass sie bestimmte nachfolgende Redezüge verlangt (co-selection).

Mit dem *SPRAC-Modell* beschreibt Linell (2009, S. 319f.) fünf Aspekte, die Konstruktionen formal erfüllen müssen.¹⁴

¹³ «A group of conversation analytic studies have focused on the therapist's formulations of the client's talk. The notion of formulation is central in conversation analysis, and it refers to instances whereby one speaker provides the gist or an upshot of another speaker's previous turn before moving on with their own contribution to the interaction. While formulations are offered as mere summaries of the previous speaker's account, they are, however, usually highly selective readings, which transform the meaning of previous talk in the directions desired by the person who utters them, while at the same time attributing this re-formulation to the initial speaker» (Georgaca & Avdi, 2009, S. 239).

¹⁴ Die Reihenfolge der Items entspricht nicht dem Kürzel SPRAC, ist aber von Linell bewusst so gehalten (S. 319).

1. *S = das Gesagte* (what is said), die lexikalische, morphosyntaktische und prosodische Struktur der Äußerung, umfasst sowohl das explizit Ausgedrückte als auch implizite Bedeutungen.
2. *R = Responsivity and Retroconstruction*: Dieser Aspekt positioniert die Konstruktion in Bezug auf die vorangegangenen Äußerungen und die Situation (responsiv, retrospektiv, retroaktiv oder retrokonstruktiv) und kann z. B. bestimmte Inhalte hervorheben oder ausblenden.
3. *P = Projektion*: Dieser Aspekt bezieht sich auf die Projektionsmöglichkeiten; die Konstruktion ermöglicht, suggeriert oder erfordert gewisse nachfolgende Konstruktionen (z. B. Antworten, Interpretationen). R und P sind selektiv und ermöglichen eine Relevanzsetzung.
4. *A = Kommunikative Aktivität*: Dieser Aspekt zeigt auf, welchen Beitrag die Konstruktion zur Ausführung einer Handlung (activity) oder für ein lokales kommunikatives Projekt leistet. Was bewirkt die Äußerung in der aktuellen Gesprächsphase?
5. *C = Die Kookkurrenz (linguistischer Ressourcen)* bezieht sich auf die charakteristischen linguistischen Merkmale, die (oft) gemeinsam mit einer spezifischen Konstruktion auftreten.

Dieses stark an der Interaktion und einer dialogischen Sicht der Kognition angelehnte Modell eröffnet neue und vielfältige Möglichkeiten der Beschreibung von Konstruktionen. Dabei stellt sich die Frage, wie die interaktionalen Merkmale der Konstruktionen abgebildet werden können. Wie kann man den dynamischen Charakter des *talk-in-interaction* (ten Have, 2007) in die Konstruktionsbeschreibungen integrieren?

Für die Lexikongestaltung müssen die beschriebenen Modelle erwogen werden, dürften jedoch schwierig umzusetzen sein. Sie führen direkt weiter zu der Frage, was nach dem Lexikon kommt, d. h. wie die interaktive, im Dialog und in der Situation «neu» konstruierte Bedeutung modelliert werden könnte, wenn sie nicht im Lexikon formalisiert werden kann (Kap. 10). Imo (2007, S. 13) spricht von einer *Interaktionalen Semantik*: Der semantische Gehalt von Phraseologismen und Konstruktionen ist zwar von der Interaktionssituation abhängig, aber Imo betont, dass Bedeutung nicht nur ad hoc und lokal entstehen kann, sondern es gibt eine *Kernbedeutung*, die im Rahmen des Konstruktionskontextes angepasst und verändert wird (siehe auch Kap. 3.3).

Wie hängen die Begriffe *Kernbedeutung* und *wörtliche Bedeutung* zusammen? Siepmann (2006) stellt die Kernbedeutung des einzelnen Wortes in

Frage und weist darauf hin, dass sie keine Relevanz hat für den Muttersprachler, d.h. die dekontextualisierte Bedeutung ist ein Artefakt und existiert nur in der linguistischen Abstraktion! Er verweist damit auf Feilke (1996, S. 128); dieser zeigt am Beispiel von verschiedenen Ausdrücken (Konstruktionen) mit scheinbar klaren Kompositionsbedeutungen, wie «Sonne», «sonnig», «scheinen», «Platz», «Plätzchen», «Lage», dass die Kompositionsbedeutung eine Illusion ist, und dass die Funktions- oder Gebrauchsbedeutung nur in der Kollokation sichtbar wird.

- (24) a. Die Sonne scheint
 b. Ein Platz an der Sonne
 c. ein sonniges Plätzchen
 d. eine sonnige Lage

So ist (24a) eine Aussage über das Wetter und nicht über die Sonne, während das Idiom (24b) eine privilegierte Situation bezeichnet,¹⁵ nicht aber einen konkreten von der Sonne beschienenen Platz wie in (24c). Beispiel (24d) dagegen ist eine Lagebezeichnung für feststehende Gebäude, kann aber nicht z. B. für Personen verwendet werden.

Gemäss Bergs und Diewald (2008b) entsprechen Konstruktionen formelhaften, mehr oder weniger fixen Ausdrücken, wie sie z. B. von Wray (2005) vorgestellt werden; einen Konsens in der Linguistik über den Platz und den Stellenwert von Konstruktionen scheint es nicht zu geben (Bergs & Diewald, 2008b, S. 1). Bergs und Diewald unterscheiden die Begriffe *construction* vs. *construct*. *Konstruktionen* sind die abstrakten Konzepte oder Muster (blueprints), die *Konstrukte* sind die konkreten Realisierungen der Konstruktionen (S. 5). Ausserdem unterscheiden sie verschiedene Ebenen von Konstruktionen, nämlich die *macro-constructions* gegenüber den *meso-constructions* und den *micro-constructions*, ein Konzept, das einleuchtend ist, um die verschiedenen Ebenen zu beschreiben (Konstruktionen im Satz, Konstruktionen in der lokalen Interaktion und Konstruktionen im grösseren Diskurskontext). Schliesslich betonen die Autoren das Ubiquitäre des Begriffs Konstruktion: «If something is not a construction, what is it?» (Bergs & Diewald, 2008b, S. 7). Dies ist auch die Antwort auf die Frage, wie umfassend Konstruktionen sind, d. h. wie viel von der Sprache letztlich ausserhalb von Konstruktionen zu verorten ist. «Nothing can be left outside, simply by definition», zu diesem Schluss kommen Östman und Fried (2005a).

¹⁵ gemäss Feilke eine historische Anspielungsbedeutung auf die afrikanischen Kolonien, «wo die Sonne immer scheint».

6.8 Grammatikalisierung

Wenn die CxG auf dem Sprachgebrauch basiert (usage-based), muss sie gleichzeitig dem Umstand Rechnung tragen, dass sich die Sprache *im Gebrauch* und *durch den Gebrauch* laufend verändert. Korpora geben einen Eindruck der Sprachverwendung in einem bestimmten Moment. Grammatische Konstruktionen können in diesem Prozess verändert werden oder auch erst daraus entstehen. Pointiert könnte man sagen, wie dies beispielsweise Haspelmath (2002) tut, dass Grammatik ein Nebenprodukt des Sprechens ist, oder «geronnener Diskurs». Redewendungen und sprachliche Äusserungen, die rekurrent gebraucht werden, verfestigen sich formelhaft und setzen damit einen Prozess der *Grammatikalisierung* in Gang. Damit einher geht eine Verminderung der lexikalischen Bedeutung (Feilke, 2007, S. 69). Im Zuge von Prozessen sprachlichen Wandels bilden sich neue Strukturen; so beschreiben Auer und Günthner (2003) die Entstehung von *Diskursmarkern* aus Matrixsätzen mit Verben des Meinens, Wissens und Sagens. Der Satz «ich glaube» wird z. B. zum Diskursmarker «glaub». Die inhaltlichen Bedeutungen der Konstruktionen verblassen, die Syntax verändert sich teilweise, wie im Beispiel des Diskursmarkers «ich mein» (Weber, 2010, S. 243). Prozesse der Grammatikalisierung und des Sprachwandels werden auch in einem Sammelband über *Konstruktionen und Sprachwandel* erörtert (Bergs & Diewald, 2008a).

6.9 Konstruktionsgrammatik: Darstellungsformalismen

Für das vorliegende Projekt ist es wichtig, die Konstruktionsgrammatiken hinsichtlich ihres praktischen Nutzens bei der Darstellung der zu untersuchenden sprachlichen Phänomene zu beurteilen. Für die Darstellung von Konstruktionen gibt es keine einheitlichen Vorstellungen; je nach Schwerpunkt, der von den verschiedenen Forschern gesetzt wird, gibt es von stark formalisierten und an gängigen Grammatikmodellen orientierten Ansätzen bis zu mehr oder weniger freien und unsystematisierten Beschreibungen aller möglichen Spielarten. Die wichtigsten Varianten sollen im Folgenden in Kürze beschrieben werden, um später eine Darstellungsform wählen zu können, die für die Abbildung von Konstruktionen im JAKOB-Lexikonprojekt optimal ist. Die gebräuchlichste Art, Konstruktionen darzustellen, ist, sie als geschachtelte Strukturen (*nested boxes*) zu zeichnen (Fried & Östman, 2005, S. 1761).

Abbildung (6.2) zeigt die Konstruktionsdarstellung mit *nested boxes* bei Fillmore (1987), wie oben erwähnt.

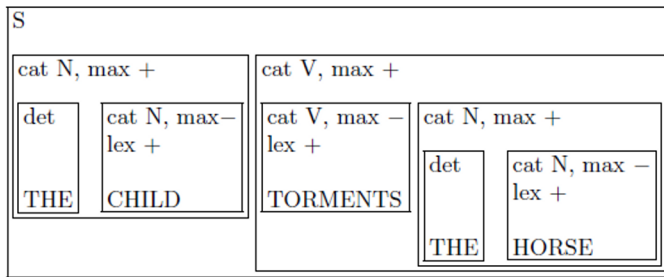


Abbildung 6.2: Notationsweise bei Fillmore (1987)

Die *spezifische Konstruktion Vergewisserungssignal* [verstehst du, verstehen Sie] wird bei Imo folgendermassen dargestellt (Imo, 2007, S. 289):¹⁶

Vergewisserung	SYNTAX	äusserungsfinal
	MORPHOLOGIE	in der Höflichkeitsform «verstehen Sie»;
	SEMANTIK	«jemanden/etwas begreifen»
	PRAGMATIK	(LEER)
	FUNKTION	Aufmerksamkeit oder Reaktion einfordern;
	SEQUENZ	(LEER)
	PROSODIE	«Frageintonation»

Abbildung (6.3) zeigt ein Beispiel der Konstruktionsdarstellung bei Goldberg (1995).

(Imo, 2007, S. 40) schlägt eine Darstellung der Konstruktionen in Wert-Attribut-Kästen vor (Tabelle 6.2). Dies ist ein intuitiv gut zugängliche Darstellungsform, gibt allerdings wenig Hilfen für eine mögliche Implementierung. Die Werte zu den einzelnen Attributen sind fakultativ, d. h. das Attribut wird nur aufgeführt, wenn es für die Konstruktion relevante Information wiedergibt. *Prosodie* wird bei Imo nur beachtet und für die Konstruktionsbeschreibung verwendet, wenn sie für die Konstruktion wichtig ist.

Auer (Auer, 2006, S. 298) verwendet ebenfalls eine einfache Merkmalsliste für die Darstellung der *Quotativ-Konstruktion mit «so»*, siehe Tabelle (6.3).

¹⁶Imo unterscheidet zwischen spezifischen, teilspezifischen und schematischen Konstruktionen.

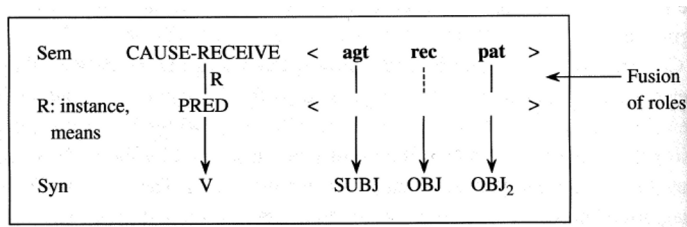


Abbildung 6.3: Notationsweise bei Goldberg (1995)

Tabelle 6.2: Konstruktionsschema (nach Imo, 2007)

	Konstruktionsname
Syntax	Einträge zur topologischen Stellung, Reihenfolge etc.
Morphologie	Morphologische Besonderheiten
Semantik	Prototypische Bedeutung
Pragmatik	Merkmale der Situation und des Diskurskontextes
Funktion	prototypische und kontextspezifische Funktionen der Konstruktion
Sequenz	Infos über sequenziellen Rahmen, über Gattung, über spezifische Interaktionsmuster
Prosodie	Prosodische Merkmale, wie Intonation, Stimmlage etc.

Zeschel (2008), der Funktionsverbgefüge untersucht, stellt die Konstruktionen in Klammernotation dar. Das Funktionsverbgefüge «den Stein ins Rollen bringen» wird bei Zeschel (als schematisierte Konstruktion) so dargestellt:

$[(NP_{agt}) NP_{thm} [in [N_{bewegungart+weise}] V]$

Stefanowitsch (2010)¹⁷ verwendet eine einfache Klammernotation, um Konstruktionen auf verschiedenen Ebenen (Wortstamm, Wort, Ausdruck, Satz) darzustellen:

- (25) Konstruktion: $[V_X\text{-er}] = \text{Drucker, Speicher etc.}$
 Konstruktion: $[\text{geb} -]$ Wortstamm, z. B. für das Verb «geben».

¹⁷Diese Beispiele stammen aus dem Handout von Anatol Stefanowitsch, IDS-Tagung Mannheim, Februar 2010.

Tabelle 6.3: Konstruktionsschema (nach Auer, 2006)

	Quotativ-Konstruktion mit «so»
Funktion/Handlungstyp:	Redewiedergabe
Gattung:	Narrativ
Sozialer Stil:	jugendsprachlich
Projektion:	Vorsatz > Nachsatz

Konstruktion: [geber] voll lexikalisiert: Geber, jemand der etwas gibt.

Konstruktion: [jmmd Saures geben] Idiom.

Konstruktion: [sich ADJ_X geben] = «den Eindruck erwecken, X zu sein».

Konstruktion: $VP[NP_{DAT}^X NP_{AKK}^Y V] = Y$ zu X transferieren.

Konstruktion: [WH VP] = Interrogativ.

Für alle diese Konstruktionen gilt, dass sich die Konstruktionsbedeutung nicht aus den Komponenten erschliessen lässt.

6.10 Fazit zur Konstruktionsgrammatik

Das Konzept der *Konstruktion* ist ein überzeugendes Modell für Lexikoneinträge und kann als Vorbild für die Lexikonentwicklung dienen. Gegenüber den vorher beschriebenen Phraseologismen hat die Konstruktion den Vorteil, dass sie von Einzelworten über Mehrwortausdrücke bis zu komplexeren satzähnlichen Gebilden jede Form annehmen kann. Als theoretisches Konzept ist die Konstruktion in vielen Bereichen der Linguistik anerkannt, aber nur selten in praktischen Anwendungen umgesetzt. Wie Konstruktionen strukturiert und dargestellt werden sollen, darüber besteht kein Konsens. Für das vorliegende Projekt eines Konstruktionslexikons soll eine *Konstruktion* einen *Lexikoneintrag* bilden und aus einer Liste von Merkmalen bestehen, nicht aus einer Kurzformel (vgl. Kapitel 9). Diese Merkmalsliste ist relativ ausführlich und besteht sowohl aus morphologischen und syntaktischen, aber auch aus semantischen und pragmatischen Attributen, für Verben ausserdem aus einem Satzmuster, das die Konstruktion auf Satzebene beschreibt (siehe folgendes Kapitel zur Corpus Pattern Analysis).

Im JAKOB-Lexikon werden nur lexikalisierte und teillexikalisierte Kon-

struktionen aufgenommen; rein schematische Konstruktionen sind nicht explizit aufgeführt, sondern nur als Muster der Instanzen, die sie repräsentieren (z. B. Lexikoneintrag «werfen» als Instanz der Transitivkonstruktion).

Ein *Problem* des Konstruktionskonzepts für das Lexikonprojekt wurde bereits im Fazit zur Phraseologie angesprochen: Lexikoneinträge als *Form-Bedeutungspaare* ergeben eine hohe Redundanz im Lexikon; bei konsequenter Durchführung des Prinzips ergeben sich viele einzelne Einträge mit minimal unterschiedlicher Form oder Bedeutung. Für das Prototypenstadium, in dem sich das Lexikon noch für einige Zeit befinden wird, kann dieses Problem jedoch vernachlässigt werden.

Ein weiteres Problem ergibt sich für Konstruktionen, die über das Satzformat hinausgehen, wie im Kapitel (6.7) als sinnvolle Erweiterung des Konstruktionsbegriffs vorgeschlagen wird. Das Konzept des JAKOB-Lexikons beruht auf der Verortung von Ausdrücken in Sätzen; wie Konstruktionen grösser als Sätze (z. B. Sprichwörter) im Lexikon abgebildet werden sollen, ist im Moment nicht klar. Für diese Fälle und generell für die Erfassung von Konstruktionen auf Diskursebene müsste eher ein separates «Diskurslexikon» geschaffen werden.

Das stark an der Interaktion und am Dialogischen orientierte Konstruktionsmodell von Linell (2009)¹⁸ zeigt neue Möglichkeiten auf und könnte als Ausgangspunkt für die weitere Entwicklung des Lexikonkonzepts dienen. Besonders die Aspekte der *Responsivität* und der *Projektion* bezüglich des Diskursablaufs müssten sorgfältig geprüft werden.

Eng mit diesen Konzepten verknüpft ist die schon mit dem Begriff der Bedeutungskonstitution aufgeworfene Frage, wie *dynamische, d. h. in der Interaktion neu kreierte Konstruktionen* in einem Lexikonmodell behandelt werden sollen. Dies ist eine der Hauptfragen, die aus der vorliegenden Arbeit am Lexikon entstanden sind; sie wurde schon in den Kapiteln zur Interaktionalen Linguistik und zur Phraseologie aufgeworfen und wird auch im Schlusskapitel eines der zentralen Themen sein. Eine mögliche Antwort, wenn auch für die Umsetzung des Lexikons nicht von unmittelbarem Nutzen, ist der Verweis auf eine *Kernbedeutung* (Imo, 2007), die in lexikalisierten Konstruktionen abgebildet wird. Diese Kernbedeutung kann mit den Anforderungen der Gesprächssituation und mit den Wünschen und Befürchtungen der Gesprächsteilnehmer durch ad hoc abgeänderte oder neu gebildete Konstruktionsbedeutungen ergänzt oder ersetzt werden.

¹⁸vgl. Kapitel (6.7).

6.11 Corpus Pattern Analysis

Corpus Pattern Analysis (CPA) ist ein neues Verfahren, um die Bedeutung von *Wörtern* (in Verbpatterns) mit den Mustern ihres Vorkommens im *normalen* Gebrauch in Verbindung zu bringen (Hanks, 2004; Hanks & Pustejovsky, 2005; Hanks, 2008a; Pustejovsky et al., 2004). Neu ist an diesem Ansatz, dass die Bedeutung strikt an den *Sprachgebrauch* gekoppelt wird. Herkömmliche Wörterbücher liefern keine hinreichenden Informationen, um verschiedene Wortbedeutungen zu disambiguieren. Die Wortbedeutung kann nicht aus vordefinierten Listen erschlossen werden, sondern ergibt sich aus der Verwendung der Wörter in bestimmten syntaktischen und phrasalen Mustern und Kontexten, d. h. jedes Wort ist mit einem oder mehreren syntagmatischen Mustern assoziiert. Wortbedeutungen entstehen in prototypischen syntagmatischen Mustern und Satzkontexten. Die Disambiguierung einzelner Wortbedeutungen erhält damit eine neue Grundlage. Ein *Verbpattern* (Muster) ist in dieser Auffassung ein semantisch motivierter und wiederholt auftretender Phraseologismus (Jezek & Hanks, 2010, S. 8).

Die Entwicklung der CPA ist kombiniert mit dem Projekt eines *Pattern Dictionary of English Verbs* (PDEV), das sich zurzeit im Aufbau befindet und online zugänglich ist.¹⁹ Das PDEV umfasst aktuell ca. 6000 englische Verben. Die *Verbpatterns* im PDEV werden aus dem BNC-Korpus²⁰ empirisch erschlossen. Die Verfügbarkeit von grossen, für elektronische Verwendung aufbereiteten Korpora, aber auch die Verfügbarkeit von geeigneten Werkzeugen und Prozeduren, um relevante und signifikante Wortmuster aus den Korpora extrahieren zu können,²¹ waren deshalb Voraussetzungen, um herausfinden zu können, dass Sprache im Gebrauch zu einem grossen Teil aus rekurrenten Mustern besteht.²²

Im *Pattern Dictionary* wird Bedeutung an minimalen Sätzen, bestehend aus einem Verb und seinen Komplementen und Ergänzungen festgemacht, nicht an einzelnen Wörtern. Verbpatterns sind Kollokationen, die um ein

¹⁹URL: <http://nlp.fi.muni.cz/projekty/cpa/> (10.9.2010)

²⁰«The British National Corpus (BNC) is a 100 million word collection of samples of written and spoken language from a wide range of sources, designed to represent a wide cross-section of current British English, both spoken and written.» Quelle: <http://www.natcorp.ox.ac.uk/> (15.11.2010)

²¹z. B. die *sketch engine*, siehe Kap. 8.1

²²«All this is possible because one of the most important findings of corpus analysis in the last 25 years has been that natural language is much more regularly patterned than many theorists assumed before corpus evidence was available» (Hanks, 2010, S. 4).

Verb angeordnet sind und deren gemeinsames Auftreten statistisch signifikant ist, d. h. nicht jede zufällige Anordnung von Einzelwörtern ist ein Muster. Verbpatterns bestehen aus den Argument- und *Valenzstrukturen* eines Verbs. Besonders für das Deutsche sind Valenzstrukturen wichtig, weil sie zusätzliche syntaktische Informationen beinhalten, wie sie durch die reichhaltigeren Morphologieinformationen des Deutschen entstehen.

Als Beispiel zitiert Hanks (2010) das englische Verb «blow»: Die zwanzig im BNC häufigsten Patterns mit dem Verb blow (davon die Hälfte Funktionsverbgefüge) machen ca. 58% der Vorkommen in «normalen» Texten aus. Jedes dieser Pattern repräsentiert eine phraseologische Bedeutung, die in der einen oder andern Form in den Köpfen der meisten Englisch Sprechenden gespeichert sein muss; andernfalls wäre Kommunikation unmöglich! Die restlichen 42% der Vorkommen repräsentieren idiomatische Ausdrücke und unübliche Verwendungsweisen. Weil Sprache ein kreatives Medium ist, gibt es kreative und manchmal auch eher bizarre phraseologische Neuschöpfungen und Erweiterungen. Idiomatische Ausdrücke sind selten, aber sie sind kognitiv salient, d. h. sie werden kaum vergessen und scheinen deshalb häufiger als sie es tatsächlich sind (S. 4). Dieser Befund wird durch die Arbeit mit den Korpora im vorliegenden Projekt bestätigt (vgl. Kap. 9.2).

Ein Hintergrund der CPA ist die Theorie der *norms and exploitations* (Hanks, 2004, in press). CPA untersucht die *normalen* Bedeutungen von Verbpatterns, daneben gibt es aber, wie im vorangehenden Beispiel ersichtlich, auch die kreative *exploitation* von bestehenden Verbpatterns. Das Erlernen einer Sprache erfordert einerseits die Kompetenz, Wörter regelhaft im wörtlichen und idiomatischen Sinn anzuwenden (*norms*), als auch die Kompetenz, kreativ neue und von diesen Regeln abweichende Verwendungsarten dieser Wörter zu erkennen oder selbst zu erfinden (*exploitations*). Hanks spricht in diesem Zusammenhang von einer *Doppelhelix-Theorie* der Sprache mit zwei komplementären Regelsystemen:

CPA does not attempt to explain all possible meaningful uses of words by means of a single monolithic theory. Instead, it proposes to explain all normal uses, leaving creative exploitations of normal uses for separate investigation. The result is a «double helix» theory of language. Human linguistic behaviour can indeed be explained as rule-governed behaviour. However, there are not one but two sets of rules. In the first place there are rules governing normal linguistic choices: standard, rather formulaic uses of linguistic patterns. These rules govern the majority of everyday uses of words. But then there are also rules governing

the exploitation of norms, for example creative metaphors and anomalous arguments. (Hanks, 2010, S. 5)

Verbpattern bestehen aus Verben und Valenzen und der Zuordnung von semantischen Typen zu den Argumenten in diesen Valenzen. Die Argumente sind *semantische Typen*, die durch *lexikalische Sets* (lexical sets) repräsentiert werden. Ein *lexical set* ist eine paradigmatische Liste von Wörtern, die an dieser Patternposition vorkommen können, wie in (26) für das englische Verb «attend» gezeigt wird.

- (26) [[Human]] attend [[Activity]]
 [[Activity]] = {meeting, conference, funeral, ceremony, course, school, seminar, dinner, reception, workshop, wedding, concert, premiere...}

Verbbedeutungen sind also verknüpft mit den semantischen Typen und lexikalischen Ausprägungen, die das Verb umgeben. Die lexikalischen Sets und die semantischen Typen sind in einer hierarchischen *Ontologie* angeordnet. Praktisch ergeben sich dadurch zwei Probleme (Jezek & Hanks, 2010, S. 8):

1. Die paradigmatischen Sets, die eine bestimmte Argumentstelle in einem bestimmten Verbpattern darstellen, lassen sich nicht zwanglos in eine einzige konzeptuelle Kategorie einordnen; sie enthalten oft Wörter, die verschiedenen semantischen Typen angehören.
2. Die Zusammensetzung dieser Wortsets ändert sich von einem Verb zum andern, d. h. es scheint keine Generalisierung über die Zugehörigkeit eines Wortes zu einem bestimmten semantischen Typ möglich. Hanks und Jezek (2008) sprechen deshalb von *shimmering lexical sets*.

Jezek und Hanks führen diese Probleme darauf zurück, dass die Ontologie nicht genügend auf realen Korpusdaten begründet ist (S. 8). Daraus folgt ein weiteres Ziel von CPA: eine Ontologie zu bauen, die diesen Anforderungen genügt (Die von CPA verwendete Ontologie wird im Kap. 7.3.3 vorgestellt). Als Alternative zu einer Ontologie wird oft eine computergestützte Cluster-Analyse vorgeschlagen (vgl. Kapitel 4.4). Aber lexikalische Sets sind offen; eine Clusteranalyse kann keine abschliessende Liste hervorbringen, der semantische Typ [[Human]] z. B. kann nicht durch eine eindeutige Lexemliste repräsentiert werden.

Im Rahmen der CPA werden verschiedene Lösungen für die oben erwähnten Probleme vorgeschlagen:

1. Nicht-kanonische lexikalische Items werden in bestimmten Kontexten in ein lexikalisches Set «gezwungen» (*coercion*). Beispielsweise werden die Nomen «Schule», «Kirche» und «Klinik» im Kontext des Verbs «besuchen» in die Kategorie [[Activity]] gezwungen, während sie in andern Kontexten zur Kategorie [[Location]] gehören.
2. Die semantische Ontologie ist nicht eine rigide Ja/Nein Struktur, sondern eine statistisch verankerte *Sammlung von kollokationalen Präferenzen*: «shimmering lexical sets». Deshalb erhält jedes Mitglied eines lexical sets statistische Informationen zum Kollokationsverhalten²³ (Jezek & Hanks, 2010, S. 12).

Lexikographen können *coercions* auf verschiedene Arten im Pattern Dictionary (PDEV) kodieren (nach Jezek & Hanks, 2010, S. 15). Der *semantische Typ* kann angepasst werden; häufig ist z. B. die Alternation *Mensch* > *Institution* (27). Alternationen können sowohl in Subjekt- als auch in Objektposition vorkommen (28).²⁴

- (27) [[Human | Institution]] hoffen
der Mensch hofft – die Organisatoren hoffen.
- (28) [[Human | Anything]] beruhigen [[Human | Body Part | Emotion]]
ein Baby beruhigen.
Schokolade beruhigt die Nerven.
der Staat will sein Gewissen beruhigen.
Beruhige dich!

Statt einer Alternation des semantischen Typs kann auch die *semantische Rolle* angepasst werden (29). Der semantische Typ ist ein *intrinsisches Attribut* eines Nomens, während die semantische Rolle durch den *Kontext* bestimmt wird. Das Beispiel zum Verb «landen» (30) demonstriert die kontextuell bestimmten Rollen «Pilot» und «Passagier». Die Adverbialergänzung ist fakultativ.

- (29) ... bringen [[Human]] {into [[Location = Jail]]} .
...versucht habe, Deinen Vater ins Gefängnis zu bringen.
- (30) [[Plane | {Human = Passenger | Pilot}]] landen
([Adv][Location])

²³vgl. *salience score* in der *sketch engine*:

<https://trac.sketchengine.co.uk/raw-attachment/wiki/SkE/DocsIndex/ske-stat.pdf>.

²⁴Beispiele aus dem Korpus *deWAC* (Beschreibung der Korpora siehe Kapitel 8.1.1).

der Pilot landete sicher auf der Piste.
 wir landeten in Zürich.
 das Flugzeug landete auf einem Feld.

Die dritte Möglichkeit besteht darin, das Pattern als *Exploitation* zu werten, wenn ein offensichtlich kreatives oder von der Norm abweichendes Pattern vorliegt. Alternationen sind reguläre Auswahlmöglichkeiten in konventionalisierten Verbpatterns, während Exploitations in der aktuellen Situation unbekannte Patterns darstellen, wie neue Metaphern (31). Dieses Beispiel stammt aus dem Korpus WIL und ist ein einzelnes Vorkommen des Verbpatterns. Im Korpus deWAC ist das Pattern ‹Dampf loswerden› nicht belegt (zum Vergleich: ‹Dampf›: 7264 Instanzen, ‹loswerden›: 7021 Instanzen, ‹los werden›: 1284).

- (31) [[Human]] loswerden {seinen Dampf}
 der Chef will seinen Dampf loswerden.

Exploitations können im Laufe der Zeit durch andere Sprachbenutzer übernommen werden und so in den normalen Sprachgebrauch übergehen (Lexikalisierung und Grammatikalisierung).

Es folgen zwei einfache Beispiele für idiomatische Verbpatterns (original englisch und deutsch).

- (32) to cry for (the) moon (Unmögliches verlangen)
 [[Human]] cry [NO OBJ] {for {moon}}
 Implicature: [[Human]] wants to have something that cannot be had.
- (33) den Stein ins Rollen bringen
 [[Human]] bringen {den Stein}{ins Rollen}
 Implikatur: [[Human]] bringt eine Sache in Gang.

6.11.1 Notation im Pattern Dictionary

Die folgenden Regeln sind aus den oben erwähnten Studien und Beispielen erschlossen, die zweite wichtige Quelle ist das öffentlich zugängliche PDEV, das aktuell ca. 6000 Verben umfasst.²⁵ Angestrebt ist nicht die Vollständigkeit der CPA-Regeln, sondern eine für das vorliegende Lexikonprojekt brauchbare und umsetzbare Version.

²⁵URL: <http://nlp.fi.muni.cz/projekty/cpa/> (18.01.2010)

- *Verbpatterns* bestehen aus dem Verb im Zentrum und sämtlichen Argumenten (dependenzgrammatisch: Ergänzungen und Angaben, inklusive Subjekt) auf der semantischen, syntaktischen oder lexikalischen Ebene.
- *Doppelte eckige Klammern* umschliessen *semantische Typen*; die Kategorien entstammen der CPA-Ontologie.
- Optionale *semantische Rollen* sind in die semantischen Typen integriert (Bsp. [[Human=Doctor]]).²⁶
- *Geschweifte Klammern* umschliessen syntaktische Konstituenten, die einen semantischen Typ enthalten können oder ein konkretes Lexem aus dem zugehörigen lexikalischen Set (Pustejovsky et al., 2004).²⁷
- *Einfache eckige Klammern* bezeichnen syntaktische Kategorien, nicht konkrete Instanzen. Z. B. bezeichnet [MOD] eine unspezifische Modifikation.
- Runde Klammern bezeichnen optionale Elemente.
- *Implikatur*: Die Bedeutung eines Patterns wird in einer Art Paraphrase beschrieben, die der Lexikograph kreieren muss. Der Begriff Implikatur stammt aus der Pragmatik und bezeichnet einen Bedeutungsaspekt, der durch die Äusserung zwar kommuniziert, aber vom Sprecher nur angedeutet anstatt explizit gesagt wird.

6.11.2 Vergleich CPA – CxG – FrameNet

Worin bestehen die Unterschiede zwischen der Copus Pattern Analysis und FrameNet? FrameNet hat ja ebenfalls Verbpatterns zum Thema.

CPA ist auf dem Hintergrund von grossen Korpora aufgebaut. Theoretische Anknüpfungspunkte sind z. B. die systemisch-funktionale Grammatik (Halliday & Matthiessen, 2004) und die Korpuslinguistik, wie sie von John Sinclair betrieben wurde. CPA analysiert Korpora Wort für Wort, oder Verb für Verb. Nach Aufnahme eines Verbs ins Patten Dictionary werden alle dazu passenden Patterns aus den zu Grunde liegenden Korpora erfasst und analysiert, womit der Prozess für dieses Verb abgeschlossen ist.

²⁶Definition bei Hanks: «Semantic type is an intrinsic property of a lexical item. Semantic role is assigned by the context» (Hanks & Pustejovsky, 2005).

²⁷Im PDEV als *Lexset* bezeichnet. «LexSet» = Lexikalisierte Instanzen zu den semantischen Typen.

FrameNet basiert dagegen auf dem Prototyp-Gedanken, die Entwicklung geht vom einzelnen Frame aus, und die Framestruktur ist bis heute sehr fragmentarisch und lückenhaft. Prototypische Frames werden für Frame entwickelt. Um das System zu vervollständigen, müssten alle möglichen Frames a priori gedacht werden (Pustejovsky et al., 2004).

6.11.3 Fazit zur Corpus Pattern Analysis

Die Ideen und Formalismen der CPA geben für das vorliegende Lexikonprojekt äusserst nützliche Hinweise und ergänzen hervorragend bestimmte Eigenarten des JAKOB-Lexikons, wie die Wichtigkeit der Verben und ihrer Valenzstrukturen, aber auch die Segmentierung der Erzähltexte in Subjekt-Prädikat-Strukturen. Besonders für das Deutsche sind Valenzstrukturen wichtig, weil sie syntaktische Informationen beinhalten, wie sie durch die reichhaltigen Morphologieinformationen des Deutschen entstehen. Ein Problem für die Anwendung von Patterns für das Deutsche könnte demgegenüber die Wortreihenfolge im Pattern sein, die für das Englische weitgehend festgelegt ist, während das Deutsche eine freiere Konstituentenreihenfolge erlaubt. In der Lexionumsetzung wird die Reihenfolge der Argumente im Verbpattern nach syntaktischen Kriterien festgelegt (Subjekt > Verb > Akkusativobjekt > Dativobjekt > Präpositionalobjekt), ist aber nicht zwingend, weil die Markierung der Abhängigkeiten und die explizite (syntaktische) Benennung der Patternelemente im neuen Merkmal <bauplan> vorgenommen wird (siehe Kapitel 9.1.3). Dieses Merkmal ermöglicht es, dass die Verbpatterns, gemessen an computerlinguistischen Massstäben der Wohlgeformtheit und Maschinenlesbarkeit, formal nicht 100%ig korrekt sein müssen, sondern sie können so gestaltet werden, dass sie für die Lexikonbenutzer transparent und einleuchtend erscheinen. Dadurch kann das Hauptgewicht bei der Darstellung auf die semantischen Eigenschaften gelegt werden; im Merkmal <pattern> sind keine syntaktischen Merkmale erfasst.

Die CPA, wie sie Hanks beschreibt, beruht zwar auf Korpora des British National Corpus, also auf schriftlichen Texten; nach den bisherigen Erfahrungen spricht aber nichts dagegen, das Konzept der Verbpatterns auch für Gesprächsdaten zu verwenden. Im JAKOB-Lexikon ist die Anwendung von Verbpatterns bereits in Angriff genommen worden (aktuell ca. 250 Verbpatterns).

6.12 Lexical Priming

Michael Hoey (2005) präsentiert seine Theorie des *lexical priming*, die in mancher Beziehung mindestens theoretisch einleuchtende, teilweise auch praktische, Lösungswege anbietet für die bereits angesprochenen Probleme der *dynamischen Entstehung und Veränderung von Bedeutung*. Die Theorie des *lexical priming* setzt das Lexikon in den Mittelpunkt des Interesses und betrachtet die Grammatik als ein Ergebnis der komplexen Lexikonstrukturen (Hoey, 2005, S. 1).

Hoey legt seine besondere Sicht auf Kollokationen dar (vgl. Kapitel 4.2.2). Er unterscheidet textuelle, statistische, psychologische und assoziative Kollokationsdefinitionen. Seine Definition der Kollokation, die für das Verständnis seiner Theorie wichtig ist, sei hier noch einmal zitiert:

So our definition of collocation is that it is a psychological association between words (rather than lemmas) up to four words apart and is evidenced by their occurrence together in corpora more often than is explicable in terms of random distribution. (Hoey, 2005, S. 5)

Hoey's Theorie hat Ähnlichkeiten mit dem psychologischen Konzept des *Priming* (semantic priming). Ein dargebotenes Wort aktiviert weitere Ausdrücke aus dem gleichen Kontext, die dann in einer nachfolgenden Aufgabe schneller abrufbar sind. Im Experiment kann z. B. gezeigt werden, dass eine vorgängige Präsentation des Wortes «body» das Erkennen des Wortes «heart» beschleunigt und erleichtert (priming word und target word). Aus diesem Konzept leitet Hoey seine Theorie ab, dass sich die Bedeutung eines Wortes durch seine Rezeptionsgeschichte entwickelt (sowohl geschrieben als auch gesprochen). Das Wort wird durch Kontext und Kotext semantisch aufgeladen. Dasselbe gilt auch über das Wort hinaus für Ausdrücke und Sätze, bei Hoey *nesting* genannt, weil die assoziativen Verbindungen nicht nur zwischen den Wörtern bestehen, sondern auch zwischen Ausdrücken und zwischen Wörtern und Ausdrücken (Hoey, 2005, S. 8). Hoey's Sichtweise hat grosse Ähnlichkeiten mit den Ideen der *Emergent Grammar* (Kapitel 3.3.5): Sprachliche Strukturen und Regeln entstehen aus dem Diskurs (emergent) und formen ebenso wieder den weiteren Diskurs, d. h. die Struktur ist ein Epiphänomen (Hopper, 1987). Grammatik ist das Ergebnis von «Routinen»; Kollokationen und Gruppierungen von Redeteilen lassen eine individuelle Grammatik entstehen (Hoey, 2005, S. 9). Priming ist nicht ein einmaliger Vorgang, sondern unterliegt ständigem Wandel, zuerst individuell, und längerfristig als Anlass für Sprachwandel.

Hoey stellt seiner Theorie *zehn Priming-Hypothesen* voran. Jedes Wort ist für den Gebrauch im Diskurs geprimed (vorbereitet, vorgespurt) durch die Erfahrung aller vorhergehenden Begegnungen des Sprechers mit diesem Wort. Primings können verknüpft werden, so dass über Priming verknüpfte Wortsequenzen selber wieder der Ausgangspunkt für weiteres Priming sind (Hoey, 2005, S. 13):

1. Every word is primed to occur with particular other words; these are its collocates.
2. Every word is primed to occur with particular semantic sets; these are its semantic associations.
3. Every word is primed to occur in association with particular pragmatic functions; these are its pragmatic associations.
4. Every word is primed to occur in (or avoid) certain grammatical positions, and to occur in (or avoid) certain grammatical functions; these are its colligations.
5. Co-hyponyms and synonyms differ with respect to their collocations, semantic associations and colligations.
6. When a word is polysemous, the collocations, semantic associations and colligations of one sense of the word differ from those of its other senses.
7. Every word is primed for use in one or more grammatical roles; these are its grammatical categories.
8. Every word is primed to participate in, or avoid, particular types of cohesive relation in a discourse; these are its textual collocations.
9. Every word is primed to occur in particular semantic relations in the discourse; these are its textual semantic associations.
10. Every word is primed to occur in, or avoid, certain positions within the discourse; these are its textual colligations.

Als wichtige Einschränkung gilt allerdings, dass alle diese Hypothesen für bestimmte Domänen und Textgenres Gültigkeit haben, aber nicht generell für die Sprache an sich (S. 13).

Ein Erklärungsansatz für die kognitive Repräsentation von Konstruktionen, der das *priming* und *nesting* von Konstruktionen beleuchtet, stammt

von Bybee (2006): Wird eine spezifische Konstruktion als Gesamtheit im mentalen Lexikon gespeichert oder werden einzelne Wörter der Konstruktion separat (distributed) gespeichert?

A phrase that is experienced only once by an adult is likely to have only a minute impact on representation compared to all of the accumulated exemplars already existing. Compare this to a young child whose experience is much more limited: each new token of experience has a greater impact on his/her representations. In addition, given a highly organized network of morphemes, words, phrases, and constructions, it will be difficult to distinguish between specific storage (storing the relatively low-frequency phrase such as *beige curtains* as a unit) and distributed storage (mapping the two words onto existing exemplars of these words) because both types of processing are occurring at the same time. Only when a sequence is repeated will access to it as a unit rather than by its parts become more efficient [...]. Thus the question of storage or nonstorage will always be a probabilistic one, based on the experience of the language user. (Bybee, 2006, S. 717)

Für das vorliegende *Lexikon- und Erzählanalyse-Projekt* sind die Ideen des *lexical priming* sehr interessant und führen direkt zur Frage, ob sich das priming, die individuelle und dynamische Bedeutungskonstitution beim Meinen und Verstehen in der Therapiesituation nachverfolgen lässt. Konstruktionen werden im Laufe der Gespräche in der Therapie wiederholt verwendet und verfestigen sich damit, werden zu *Formeln* für interaktionale und therapeutische Prozesse. Das Modell von Bybee ergänzt die Sicht auf die *Emergenz* um einen neurophysiologischen Aspekt.

Eine systematische Erhebung von Lexical Priming-Effekten in den verwendeten Korpora ist im vorliegenden Rahmen kaum zu leisten; immerhin gibt es im grössten Korpus der Therapiegespräche (WIL) Anzeichen dafür, dass sich mit der häufigen Verwendung bestimmter Ausdrücke und Phrasen durch die Klientin oder durch den Therapeuten ein gemeinsamer Erfahrungshintergrund aufbaut, in dem sich gemeinsame lokale Bedeutungen im Sinne eines Priming entwickeln. Ein Beispiel dafür sind Konstruktionen mit dem Bestandteil «Zeug», die sich im Laufe der Jahre als eine gemeinsame, wenn auch vage Begrifflichkeit entwickeln.

In der praktischen Analyse der Korpusdaten eher nachzuverfolgen ist die bereits erwähnte Theorie der *norms and exploitations* (Hanks, in press).

- (34) Dampf ablassen (deWAC: n=412)
- (35) Frust ablassen (deWAC: n=114)
- (36) seinen Dampf loswerden (deWAC: n=0, WIL: n=1))

Die Beispiele (34 - 36) zeigen auf Grund der Häufigkeit des Vorkommens der Konstruktionen in den Korpora gut auf, welche Ausdrücke als *norm* und welche als *exploitation* angesehen werden können. Während (34) und (35) als Ausdrücke aus dem normalen Sprachgebrauch angesehen werden können, handelt es sich in (36) um kreativen Sprachgebrauch.

Diese Priming-Prozesse können auch in einer gewissen Analogie zu den Statements der *narrativen Psychologie* gesehen werden, in denen es darum geht, ein neues *Lebensnarrativ* zu konstruieren; Patient und Therapeut konstruieren in gemeinsamer Formulierung ein neues Lebensnarrativ.

6.12.1 Fazit zum Kapitel Lexical Priming

Die zehn Priming-Hypothesen von Hoey (2005) ergeben einen informativen Katalog von Verknüpfungsmöglichkeiten von Wörtern und Ausdrücken auf den verschiedenen Ebenen, von der simplen Kollokation bis zu kontextuellen Präferenzen bezüglich Kohäsion, semantischen Ähnlichkeiten und Positionen im Text. Alle diese Eigenschaften des Priming können implizit als Kriterien für die Charakterisierung von Lexikoneinträgen verwendet werden, erfordern aber keine zusätzlichen Lexikonmerkmale. Die Priming-Hypothesen beleuchten vielmehr die kognitiven Hintergründe der *Verknüpfung* von lexikalischen Items, wie sie mit Hilfe des Merkmals *cross reference* im Lexikon abgebildet werden können (siehe Kapitel 9.1.2).

Letztlich liefern die Überlegungen zu den mannigfachen Verknüpfungsmöglichkeiten von lexikalischen Items zu geprägten Einheiten in erster Linie Erklärungsmodelle für die Speicherungs- und Kombinationsmöglichkeiten kognitiver Art, aber weniger konkrete Implementierungshilfen. Die *cross references* im Lexikon können zwar ansatzweise auf dem Hintergrund der Priming-Hypothesen modelliert werden, da aber das Priming dynamisch und individuell abläuft, können immer nur Teilaspekte erfasst werden. Weitere Bemerkungen zu *cross references* sind im Schlussteil zu finden (Kapitel 10.3.1).

7 Lexikon und Lexikographie: Konzepte

Im Lexikon fließen die Erkenntnisse aus den bisher vorgestellten verschiedenen Disziplinen zusammen: Narrativik, Gesprächsforschung, Phraseologie und Kollokationsforschung, Konstruktionsgrammatik, Framekonzepte und kognitive Schemata verbinden sich mit lexikographischen Konzepten. «Der unverzichtbare Kern jedes Sprachverarbeitungssystems ist ein Lexikon» (Gibbon, 2004, in Carstensen et al., 2004, S. 432). Im folgenden Kapitel werden die bisher gesammelten Ideen in Bezug zu lexikographischen Konzepten und zum konkreten Projekt des *Konstruktionslexikons* gebracht. Als Ausgangspunkt wird zuerst das bisherige JAKOB-Lexikon vorgestellt, wie es vor Projektbeginn verwendet wurde. Anschliessend wird eine Auswahl bestehender elektronischer lexikalischer Ressourcen vorgestellt, die verschiedene Wege der Darstellung lexikalischer Daten beschritten haben und die die Entwicklung des Lexikonprojekts beeinflusst haben. Aus der kritischen Auseinandersetzung mit diesen sehr unterschiedlich aufgebauten Lexika entstand das Bedürfnis, die hinter den verschiedenen Ansätzen stehenden *Kategorien* und *Ontologien* semantischer und pragmatischer Art zu erkunden, um eine passende Lösung für das eigene Projekt zu finden. Die Auswahl umfasst sowohl Kategorisierungsvorschläge für Gesprächscharakteristika als auch semantische Typologien. Schliesslich beendet ein kleiner Ausblick auf die lexikologische Bedeutung der Grundlagenkonzepte von *Nominalität* und *Verbalität* sowie auf die mögliche Bedeutung von *Konnotationen* für Lexikoneinträge das vorliegende Kapitel.

Inventare von Wörtern, Redewendungen und ihren Eigenschaften haben von jeher den Menschen fasziniert und sind Ausdruck seines enzyklopädischen Dranges. Übliche semasiologische Wörterbücher gehen von einem Wort aus und listen dessen Bedeutungen auf.¹ «Der deutsche Wortschatz nach Sachgruppen», *Dornseiff* genannt (Dornseiff, 2004) ist ein onomasio-logisches Wörterbuch, d.h. die Wörter sind gruppiert um ein System von *Sachgruppen*.

¹ engl.: sense-enumerative lexicons.

Ein Vorreiter für eine neue Sichtweise auf das Lexikon und auf die Lexikographie war *John Sinclair*. Er erschloss neue Möglichkeiten der Wörterbuchkonzeption, indem er Kollokationen und Phraseologismen als wichtige Bausteine des Lexikons erkannte; und dies auf der Grundlage von *realen Korpusdaten*, nicht von (spekulativer) Introspektion. Ermöglicht wurde dies auch durch die in dieser Zeit entstehenden elektronischen Textkorpora (Hanks, 2008b).

Als Beispiel einer elektronisch aufbereiteten lexikalischen Ressource soll das DWDS² erwähnt werden. Fellbaum, Geyken, Herold, Koerner und Neumann (2006) stellen das Design und die Entwicklung einer grossen lexikalischen Ressource für die Erfassung von deutschen idiomatischen verbalen Phraseologismen und Funktionsverbgefügen innerhalb des DWDS vor. Die erfassten Phraseologismen ermöglichen detaillierte linguistische Analysen auf der Basis der dem Projekt zugrunde liegenden Korpora (S. 357).

Schliesslich ist als eine Lexikonausprägung auch das *mentale Lexikon* zu nennen, nämlich «das neuronal basierte, internalisierte Wissen von Sprechern über Wörter und ihre Eigenschaften als Gegenstand psycho- und neurolinguistischer Forschung» (Engelberg & Lemnitzer, 2009, S. 7).

Zuerst soll nun das bestehende JAKOB-Lexikon kurz vorgestellt werden, das seinen Ursprung in der *Erzählanalyse JAKOB* hat und dort als Kodierhilfe entwickelt wurde.

7.1 Das JAKOB-Lexikon

Das dem vorliegenden Projekt zu Grunde liegende JAKOB-Lexikon ist als elektronisches Wörterbuch aufgebaut und dient als Grundlage für die *Kodierung* der mit der Erzählanalyse JAKOB aufbereiteten Erzählungen. Das Lexikon beinhaltet die Zuordnung der JAKOB-Codes zu den einzelnen Wörterbucheinträgen. Aktuell umfasst das Lexikon ca. 7000 Einträge (November 2010).

Die verwendeten Codes sind im JAKOB-Kodiersystem festgelegt (‹Dramaturgische Kodierung›), das in den Manualen zu JAKOB ausführlich beschrieben ist (Arboleda et al., 2010; Boothe, 2002; Boothe et al., 2002). Durch die lexikalische Wortwahl gestaltet der Erzähler die Dramaturgie und das psychodynamische Potenzial der Erzählung entscheidend, stellt damit eine charakteristische szenische Struktur dar und vermittelt Hinweise auf psychisches Geschehen und Konflikte. Besonderes Gewicht erhalten in diesem Zusammenhang die *Aktionen* des Erzählers, also die Verben und

²Digitales Wörterbuch der Deutschen Sprache: <http://www.dwds.de/>

Verbausdrücke. Sie geben Aufschluss über das Erlebens- und Verhaltensrepertoire des Erzählers sowie in die Dynamik der Wünsche, Ängste und Abwehrmechanismen.

Das JAKOB-Lexikon ist in der *Internetapplikation* der Erzählanalyse JAKOB³ integriert und als MySQL-Datenbank implementiert. Die Abfrage des Lexikons erfolgt im Analyseschritt der Kodierung durch die Applikation selber; das Storyvokabular wird mit Hilfe des JAKOB-Lexikons automatisch kodiert. Daneben kann das Lexikon auch als eigenständige Applikation verwendet werden, es stehen verschiedene Such- und Abfragemöglichkeiten zur Verfügung (Abfrage nach Worteintrag, Abfrage nach JAKOB-Codes, Suche nach Codedefinitionen etc.).

7.1.1 Aufbau des bisherigen JAKOB-Lexikons

Es folgt eine kurze Beschreibung des bisherigen Lexikons als Übergang von der im Kapitel zum Narrativ dargestellten Erzählanalyse JAKOB zum aktuellen Lexikon-Projekt und seinen potentiellen Fortsetzungen und Ergänzungen. Nach einer ersten Planung und Implementierung des JAKOB-Lexikons als Microsoft-Access Datenbank (Luder, 1999) konnte im Jahr 2004 eine internetbasierte Version in Betrieb genommen werden. Von Anfang an stand fest, dass das JAKOB-Lexikon als *Stammformenlexikon* ausgelegt werden soll und nicht als Vollformenlexikon, was wesentlich umfangreicher und für Wartung und Ergänzung komplizierter wäre. Ein Eintrag bestand aus einem Wort, Mehrworteinträge waren nicht möglich. Bereits in dieser Erstellungsphase war klar, dass ein Lexikoneintrag aus einem Wort, kombiniert mit einer bestimmten Bedeutung bestehen soll, d.h. es gibt mehrere gleich lautende Lexikoneinträge. Zum Zweck der Disambiguierung dieser Einträge wurde die Lexikonstruktur um die Merkmale <Subkategorisierung>, <Features>, <Domains> und <Ergänzung> erweitert.

Die Eigenschaft *Subkategorisierung* bezeichnete syntaktische Merkmale eines Verbeintrages und bestand aus dem *Satzmuster* (nach Wahrig, 2007), mit dem dieses Verb verwendet wird. Das im Lexikon so genannte Merkmal *Features* bezeichnete semantische Eigenschaften und Selektionsrestriktionen der Lexikoneinträge. Diese Merkmale dienten dazu, gleich lautende Lexikonausdrücke zu disambiguieren. Der praktische Verwendungszweck der *Features* in der JAKOB-Applikation war die Anaphernresolution, d.h. die Referenzierung der Akteure und ihrer Pronomen in den Erzählungen. Das Merkmal *Domain* bezeichnete das *Thema* oder den *Kontext*, der dem Ein-

³URL: <http://www.jakob.uzh.ch/lexikon/>

trag zugewiesen wurde. Während des automatischen Kodierens wurden die aufgetretenen Domains in einer Liste verwaltet und die Lexikoneinträge passend zu den letzten Positionen in der Liste gesucht. Als weitere Möglichkeit der Disambiguierung gleich lautender Lexikoneinträge mit unterschiedlicher Bedeutung wurde das Merkmal *Ergänzung* eingeführt, das es ermöglichte, die Stammform mit einem zusätzlichen Ausdruck zu ergänzen, der allerdings wörtlich im gleichen Segment vorkommen musste. Schließlich enthielt das Merkmal *JAKOB-Code* die passende Kodierung für den Lexikoneintrag, mit besonderem Gewicht auf den Verben.

Im Jahr 2008 wurde eine umfassende Neukonzeption des JAKOB-Lexikons auf der Grundlage der OLIF-Lexikonstruktur vorgenommen und implementiert (siehe Kap. 9.1); wesentlichste Neuerung war dabei die Erweiterung mit Mehrworteinträgen. Damit war die technische Grundlage gelegt für die im vorliegenden Projekt beschriebene Weiterentwicklung des Lexikons (siehe Kapitel 9.1).

Das *JAKOB-Kategoriensystem* wird in den Grundlagen zur Erzählanalyse JAKOB und in den Manualen beschrieben (Arboleda et al., 2010; Boothe, 1989, 2002, 2004; Boothe et al., 2002) (siehe auch Kap. 2.3.4).

7.2 Bestehende Mehrwort-Lexika

Im Folgenden soll eine Auswahl von (elektronisch verfügbaren) *lexikalischen Ressourcen* vorgestellt werden, die alle in der einen oder andern Hinsicht mit dem vorliegenden Lexikonprojekt in Verbindung gebracht werden können, indem sie als Vorbilder gewisse Standards dokumentieren oder indem sie Problemlösungsmöglichkeiten aufzeigen. Die Auflistung ist nicht systematisch, sondern repräsentiert eine willkürliche Auswahl und Reihenfolge. Als Einstieg ein Zitat von Christiane Fellbaum, die zum Projekt DWDS schreibt:

Ein digitales Wörterbuch stellt eine fantastische Alternative zum traditionellen Wörterbuch dar und eröffnet neue Möglichkeiten, das Lexikon einer Sprache abzubilden. Traditionelle Papierwörterbücher müssen gezwungenermaßen alphabetisch organisiert sein. Aber diese Organisation, obschon praktisch, ist wissenschaftlich recht uninteressant. Einfache psychologische Experimente zeigen, dass Menschen Wörter, und die dahinter stehenden Konzepte, hauptsächlich semantische Beziehungen, also Bedeutungsprinzipien, zur Organisation ihres <geistigen Lexikons>

benutzen. Ein digitales Wörterbuch lässt eine psychologisch realistische Darstellung des Wortschatzes zu. Daneben kann die orthographische Ordnung natürlich beibehalten werden. (Christiane Fellbaum: Potsdamer Universitätszeitung, Januar 2003)⁴

7.2.1 WordNet

WordNet ist ein elektronisch verfügbares lexikalisch-semantisches *Wortnetz* (Wortschatz) und bildet die häufigsten und wichtigsten Wörter der englischen Sprache und ihre bedeutungstragenden Beziehungen zu anderen Wörtern der Sprache ab. Mehrere Hypothesen waren von Anfang an begleitend für das Projekt (Fellbaum, 1998b, 1998a):

- *Separability hypothesis*: Die lexikalische Seite der Sprache kann bis zu einem gewissen Grade isoliert vom Diskurs betrachtet werden. Auf der Ebene der Wörter können substantielle Erkenntnisse zur Sprachmodellierung und -verwendung gewonnen werden. Gegen diese isolierte Betrachtungsweise gibt es aber viele und gewichtige Einwände.
- *Patterning hypothesis*: Der Mensch kann all das lexikalische Wissen, das zur erfolgreichen sprachlichen Verständigung nötig ist, nicht präsent haben, ohne auf systematische Strukturen und Muster zurückzugreifen, die dieses lexikalische Wissen ordnen und die (semantischen) Verbindungen zwischen den Einzelkomponenten herstellen.
- *Comprehensiveness hypothesis*: Sofern computerlinguistische Werkzeuge je ein gewisses Sprachverstehen erreichen sollten, dann müssten sie einen mächtigen Vorrat an lexikalischem Wissen zur Verfügung haben, wie dies der Mensch auch braucht.

Zentral für den Aufbau von WordNet ist das Konzept des *Synsets* (synonym set), das eine hierarchisch geordnete Gruppe von Wörtern zum gleichen Konzept beinhaltet. Ein Wort bildet den Konzeptknoten; von hier aus führen semantische Verknüpfungen zu Oberbegriffen, Unterbegriffen und zu verwandten Konzepten (Relationen). Für jeden Worteintrag gibt WordNet eine Anzahl Synonyme, die zum gleichen Synset gehören, mit einem Kommentar, der sowohl eine kurze Definition beinhaltet als auch Beispielsätze.

WordNet blickt im Kontext elektronischer Sprachressourcen auf eine lange Tradition zurück, eine erste Version wurde bereits 1985 vorgestellt. Die

⁴<http://www.uni-potsdam.de/portal/jan03/computer.htm> (12.11.2010)

Grundeinheit ist das Wort; allerdings enthält WordNet auch zusammengesetzte Ausdrücke, Kollokationen und idiomatische Phrasen, aber die Grundeinheit ist das Wort, WordNet enthält keine konzeptuellen und organisatorischen Einheiten, die über das Synset hinausgehen, wie z. B. Scripts oder Frames. Der Aufbau ist einem Thesaurus ähnlich; ein Synset umfasst alle Wörter, die ein bestimmtes Konzept ausdrücken und ist selber ein lexikalischer Oberbegriff dieser Wörter oder ein semantisches Konzept. *Relationen* bestehen auf der lexikalischen Ebene aus Verknüpfungen mit semantisch ähnlichen Wortformen (z. B. Synonymie oder Antonymie), auf der semantischen Ebene zwischen verschiedenen Konzepten, Hyponymie oder Hyperonymie; vgl. die umfangreiche Liste der Verknüpfungstypen in der OLIF-Lexikoneigenschaft *cross reference*, Kap. 9.1.2).⁵

Die Wortnetze sind für die syntaktischen Kategorien Nomen, Verben, Adjektive und Adverbien getrennt angelegt und umfassen nur *lexikalisches Wissen*. Das Verständnis für die Bedeutung und den Gebrauch eines Wortes erfordert aber neben dem lexikalischen Wissen vor allem auch *Weltwissen* (enzyklopädisches Wissen). Das englische WordNet ist in viele Sprachen übersetzt und in zahlreichen Studien beschrieben und angewandt worden. Die riesige WordNet-Ontologie hat jedoch oft nicht den erwarteten praktischen Wert, weil die Synsets eher intuitiv und im Hinblick auf eine praktikabel darstellbare Ontologie aufgebaut sind und den beobachtbaren Sprachgebrauch ausser Acht lassen (vgl. Hanks & Pustejovsky, 2005, S. 3).

Das JAKOB-Lexikon ist stark «verblastig» und enthält nur eine eher dürftige Repräsentation der *Nomen*. Die Verwendung einer Hierarchie von Synsets könnte die geordnete Darstellung der Nomen stark verbessern. Durch die Ergänzung der Nomen-Einträge durch semantische Typen aus einer passenden Ontologie könnte jedes Nomen in ein oder mehrere passende lexikalische Sets eingeordnet werden.

7.2.2 GermaNet

Das deutsche Wortnetz *GermaNet* hält sich an die Strukturen und an das Datenbankformat des englischen Princeton WordNet, setzt aber eigene Schwerpunkte:

GermaNet and WordNet share the same technology: the data base format and the software code to build and access the data

⁵Ein umfangreiches *Glossar* der in WordNet verwendeten Begriffe ist online zugänglich: <http://wordnet.princeton.edu/man/wngloss.7WN.html> (15.11.2010)

base are the same. Some minor modifications had to be applied to the software to reflect peculiarities of GermaNet and the German language. (Quelle: GermaNet-Homepage)⁶

GermaNet umfasst Nomen, Verben und Adjektive in Synsets mit «gleichen» oder ähnlichen Wortbedeutungen (siehe WordNet). Germanet umfasst aktuell⁷ ca. 61'000 Synsets mit 84'000 lexikalischen Einheiten (Verben: 12'000, Nomen: 64'000, Adjektive: 8100); davon nur sehr wenige Mehrwortausdrücke. *Lexikalische Relationen* bezeichnen Beziehungen zwischen Wortbedeutungen, *konzeptuelle Relationen* bilden Beziehungen zwischen den Synsets ab, die für alle Synset-Mitglieder gelten. Um die Hierarchie besser zu strukturieren und um Kreuzklassifikation zu vermeiden, wurden «künstliche» Konzepte⁸ eingeführt (Kunze & Lemnitzer, 2007). Die *Verbklassifikation* in GermaNet lehnt sich an die Klassifikation von Levin an (Levin, 1993).

GermaNet wurde auch ins *EuroWordNet* integriert, ein polylinguales Wortnetz nach dem Vorbild des WordNet für acht europäische Sprachen.

7.2.3 FrameNet

Das *Berkeley FrameNet Projekt* ist eine englischsprachige lexikalische Resource, die auf der *Framesemantik* basiert und deren Struktur aus Korpusvidenzen aufgebaut wird (Ruppenhofer, Ellsworth, Petruck, Johnson & Scheffczyk, 2010). Die Theorie der Framesemantik wird von Charles Fillmore und Mitarbeitern seit mehr als 30 Jahren kontinuierlich entwickelt (Boas, 2005; Fillmore, 1968; Fillmore, Johnson & Petruck, 2003; Fillmore, 2003). Semantische Frames bilden das Hintergrundwissen (Kontext), auf dessen Hintergrund die Bedeutung der einzelnen Lexeme interpretiert wird.

Die FrameNet-Datenbank umfasst aktuell mehr als 10'000 Lexikoneinträge in ca. 1000 semantischen Frames (Baker, Fillmore & Cronin, 2003; Ruppenhofer et al., 2010)⁹. Ein *Lexikoneintrag* (*lexical unit*, LU) ist eine Kombination eines Wortes mit einer Bedeutung. Jede einzelne Bedeutung eines polysemen Wortes gehört in der Regel zu einem separaten Frame. *Semantische Frames* sind Konzeptstrukturen, die Situationen oder Ereignisse beschreiben. Frames werden durch bestimmte Ausdrücke abgeru-

⁶URL: http://www.sfs.uni-tuebingen.de/GermaNet/germanet_structure.shtml (15.11.2010)

⁷URL: <http://www.sfs.uni-tuebingen.de/lsd/index.shtml> (August 2010)

⁸z. B. die Konzepte «?abstammender Mensch» vs. «?ausgebildeter Mensch», um die Nomen «Adliger», «Bürgerlicher», «Fachkraft» und «Akademiker» einordnen zu können.

⁹URL: <http://framenet.icsi.berkeley.edu/> (15.11.2010)

fen und halten dann ein Inventar von *Frame-Elementen* bereit, semantischen Rollen, die für den betreffenden Handlungskontext charakteristisch sind. Im einfachsten Fall wird der Frame durch ein Verb initiiert, und die Frame-Elemente sind die vom Verb abhängigen Argumente, Objekte und Ergänzungen (Prädikat-Argumentstruktur, Kernrollen, periphere Rollen und extra-thematische Rollen). Frames beschreiben Strukturen mit dem Verb im Zentrum.

Die Frames sind über verschiedene *Relationen* zu einem Netzwerk verknüpft. Zwischen einem übergeordneten und einem untergeordneten Frame sind folgende Verknüpfungen möglich:

- *IS-A relation*: Der child-frame ist ein Subtypus des parent-frame.
- *Using*: Der child frame verlangt einen parent-frame als Hintergrund.
- *Subframe*: Der child frame ist in einen übergeordneten Frame eingebettet.
- *Perspective on*: child-frame with particular perspective on an unperspectivized parent-frame.

Einen Vergleich mit WordNet machen Ruppenhofer et al. (2010). Die für den Kontext des vorliegenden Projektes wichtigsten Punkte sind:

- FrameNet-Lexikoneinträge sind durch ausführliche *Definitionen* charakterisiert, WordNet beinhaltet eine knappe Paraphrase und fokussiert auf das Synset.
- Die Beispiele zu den Frames sind aus realen Sprachkorpora entnommen und nicht von Lexikographen konstruiert. Damit repräsentiert FrameNet den empirischen Sprachgebrauch.
- FrameNet: Die Sicht auf das Lexikon ist *frame-zentriert* und nicht auf das einzelne Lemma zentriert.
- Jede LU ist mit einem *semantischen Frame* verbunden, und dadurch auch mit den andern Ausdrücken, die durch diesen Frame repräsentiert werden.
- FrameNet besitzt, wie WordNet, ein System von Verknüpfungen zwischen den Frames.
- FrameNet ist weniger als Ontologie brauchbar als WordNet.

Ein Beispiel: Das polyseme Lemma «backen» (bake) ist 3 verschiedenen Frames zugeordnet, d.h. das Lexikon enthält 3 verschiedene *Lexical Units* (bake.v) mit je unterschiedlichen Definitionen (siehe Beispiel 37). Lexical Units können auch Mehrwortausdrücke, zusammengesetzte Wörter oder idiomatische Phrasen sein. FrameNet analysiert aber die interne Struktur dieser Mehrwort-LU's nicht weiter.

- (37) a. bake.v - Frame: Absorb_heat (to be cooked by dry heat, especially in an oven and as food)
 b. bake.v - Frame: Apply_heat (cook by dry heat without direct exposure to a flame, typically in an oven)
 c. bake.v - Frame: Cooking_creation (cook food by dry heat without direct exposure to a flame, typically in an oven)

In der praktischen Arbeit am Aufbau des Lexikons hat sich gezeigt, dass das *Konzept der Frames* zwar theoretisch sehr einleuchtend ist und dass die Verwendung von Frames zur semantischen Strukturierung des Sprachwissens grosse Erkenntnisse verspricht, dass aber der praktische Nutzen eher gering ist. Erstens ist es aufwendig, die Lexikon-Items einer passenden Framestruktur zuzuordnen, noch aufwendiger wäre es die komplexe Framestruktur für die Analyse zu implementieren. Der Nutzen liegt also eher auf einer zusätzlichen deskriptiven Ebene der lexikalischen Beschreibung.

Der konzeptuelle Hintergrund von FrameNet ist in diversen Frametheorien zu finden, die auf grosses theoretisches Interesse stiessen, aber ausser im FrameNet-Projekt kaum praktisch umgesetzt wurden. Eine ausführliche Darstellung der Entstehung und der verschiedenen Ausprägungen der Frametheorien findet sich bei Konerding (1993):

Die Idee, dass Wissen im menschlichen Gedächtnis in Form komplex strukturierter Gesamtheiten vorhanden sein soll, wurde sehr schnell zu einem akzeptierten und populären Standpunkt in Psychologie, Künstlicher Intelligenz und Linguistik. Die Diskussion, ausgelöst durch Minskys Beitrag von 1975, blieb jedoch weitgehend auf theoretisch spekulative und allgemein konzeptionelle Überlegungen beschränkt. (S. 1)

7.2.4 Das UWV-Analysemodell

Beim UWV-Analysemodell handelt es sich um ein prototypisches Modell zur Darstellung von Phraseologismen (Stand 2010), hier *Usuelle Wortver-*

bindungen (UWV) genannt (Steyer & Brunner, 2009). Diese werden in grossen Korpora auf Grund häufigen Vorkommens identifiziert, qualitativ ausgewertet, zu Worverbindungenmustern zusammengestellt und in hierarchischen Strukturen abgebildet. Dieses Vorgehen liefert neue Informationen zu häufig vorkommenden, mehr oder weniger stereotypen Wortverbindungen, zu typischen Gebrauchsmustern von Wörtern und Ausdrücken und ihrer pragmatischen Relevanz. Die entstehenden *Wortbindungsfelder* werden elektronisch aufbereitet und online publiziert.¹⁰ Das Vorgehen und die Ziele werden wie folgt beschrieben:

Der Erkenntnisfortschritt unserer Methodik beruht auf der Identifikation bisher verdeckter Strukturen und Phänomene. Die automatisch ermittelten Daten liefern dabei erste starke Evidenzen. Jedoch greifen wir auf jeder Analyseebene linguistisch-interpretierend ein. Das bedeutet, dass alle Ergebnisse, die wir darstellen, vor dem Hintergrund unserer subjektiven Interpretation zu bewerten sind. Dennoch ist unser Vorgehen – unter systematischer Ausnutzung automatischer Verfahren – sehr viel strenger daten- und oberflächenorientiert als das bei den meisten vergleichbaren lexikologischen und lexikografischen Projekten der Fall ist. Daher können wir gesichere Aussagen über den tatsächlichen Sprachgebrauch treffen. (Steyer & Brunner, 2009, S. 3)

Das UWV-Analysemodell beinhaltet vier aufeinander aufbauende Interpretationsstufen:

1. Die Ebene der konkreten Oberflächenmuster in den KWIC-Zeilen (Suchmuster).
2. Die Ebene der *lexikalischen Realisierungen* (LR): typische Realisierungen einer Wortverbindung in den Korpustexten.
3. Die Ebene der *Wortverbindungen*: übersummativ Gebrauchsbedeutung. Die LR müssen einem gemeinsamen semantisch-pragmatischen Konzept zuzuordnen sein.
4. Die Ebene der *Wortbindungsmuster* (S. 10).

Eine LR-Gruppe ist die Gesamtmenge aller möglichen Realisierungsformen einer Wortverbindung. Es gibt mehrere Typen lexikalischer Realisierungen, von der Kernvariante (Bsp. «aus persönlichen Gründen») bis zu

¹⁰URL: <http://wvonline.ids-mannheim.de/> (28.8.2010)

einer *Kontext-LR*, die Informationen über Verwendungszusammenhänge, Bedeutung und Funktion der Wortverbindung in spezifischen Kommunikationssituationen enthält (Kontextmarker). So ist z. B. «aus persönlichen Gründen» oft in Zusammenhang mit den Kontextmarkern «Amt», «Funktion» oder «Posten» zu finden (S. 24f). Die lexikalischen Realisierungen werden anschliessend zu Wortverbindungen zusammengefasst. Die Wortverbindung entspricht meist der Kern-LR, weil diese den prototypischen Kern abbildet (S. 33). *Wortbindungsmuster* sind schliesslich die weiter abstrahierten Kategorien, die aus den Wortverbindungen gewonnen werden. Das folgende Zitat fasst zentrale Aspekte der Modellierung von Meinen und Verstehen prägnant zusammen und soll deshalb ausführlich wiedergegeben werden.

Die Ebene <Wortbindungsmuster>, [...] bildet [sie] kognitive Mechanismen des Sprechens und Verstehens ab, und zwar das Phänomen der dynamischen Aktualisierung von konkreten Gebrauchsbedeutungen einer Wortverbindung und dem zu Grunde liegenden Muster in einer Äusserungssituation. Die Fähigkeit von Sprechern, konkrete Äusserungen vor der Folie eines Musters höherer Ordnung zu verstehen, wird in der Phraseologie vor allem als Erklärung für das Verstehen von sprachspielerischen Modifikationen herangezogen. Sie scheint aber sehr viel weitreichender für das Funktionieren von Kommunikation zu sein und eben nicht nur im engen Sinne für den Gebrauch von Phraseologismen zu gelten. Ein Beispiel ist die Wortverbindung «aus finanziellen Gründen». Wird diese Wortverbindung in einer konkreten Äusserungssituation eingesetzt, können die Hörer natürlich ihre Bedeutung erschließen, nämlich dass die Ursachen/Motive für einen Sachverhalt finanzieller und nicht etwa technischer oder politischer Natur sind oder im privaten Bereich liegen. Gleichzeitig verstehen sie jedoch die pragmatische Funktion mit: Es wird im Vagen gelassen, welcher Art die finanziellen Gründe und was die genauen Umstände sind. Die Sprecher können diese Vagheitsfunktion sowohl einsetzen als auch decodieren, weil sie viele Verwendungsfälle ähnlicher Art bereits als verfestigte Muster im Kopf haben, und zwar als restringiertes Muster «aus X Gründen». Die Restriktion liegt im rekurrenten Gebrauch begründet und bezieht sich zum einen auf den Referenzbereich der Adjektive (domänenbezogen), wobei die einzelnen lexikalischen Vertreter frei austauschbar sind, zum anderen

auf die rekurrente Verwendung der Pluralform. (Steyer & Bruner, 2009, S. 35)

7.2.5 Idiomatisierte verbale Phraseologismen (IVP)

Unter dem Titel «Über Bärendienste und erstaunte Bauklötze» präsentiert Söhn (2006) einen Katalog idiomatisierter verbaler Phraseologismen, beschränkt auf Idiome ohne freie Lesart im Deutschen. Die Datenbank umfasst 190 Einträge (Stand 2005).¹¹ Söhn sucht mit seiner Beschreibung und Klassifizierung nach lexikalischen Kriterien, um verbale Phraseologismen adäquat und kohärent zu beschreiben. Er unterscheidet teildiomatische und vollidiomatische Phraseologismen, die charakteristische Eigenschaften und Anomalien in Bezug auf Morphologie, Syntax und Lexik aufweisen (z. B. «eingefrorene Eigenschaften» wie Numerus, Tempus oder Modus). *Syntaktische Anomalien* sind z. B. die Verwendung des Dativ statt des Akkusativs in der Redewendung «etwas kostet jmdm das Leben» oder spezielle Valenzmuster, wie in der Wendung «jmdn Lügen strafen» (Söhn, 2006, S. 16). Die Klassifikationskriterien der verbalen Phraseologismen (IVP) umfassen die in der Tabelle (7.1) dargestellten Merkmale (am Beispiel des Idioms «Dampf ablassen»).

Tabelle 7.1: Eigenschaften der IVP (nach Söhn, 2006)

Merkmal	Beispiel
IVP (Lemma):	«Dampf ablassen»
Paraphrase:	Druck abbauen, s. Ärger Luft machen
Marker:	keiner
Metaphorisch?	ja
Metonymisch?	nein
Motiviert?	ja
Dekomponierbar?	ja; Dampf-ÄRGER, ablassen-ABBAUEN
Unikales Element?	nein
Varianten:	Keine
Valenz:	intern: objNP; extern: subjNP
Restriktionen:	(siehe nachfolgende Merkmale)
referenziell:	nein;
passivierbar:	ja (aber oft in wörtlicher Lesart...)
Adjunkte?	nein;
Nicht-Negation?	ja;
Kein-Negation?	nein

¹¹ URL: <http://www.soehn.net/ivp/korpus.htm> (15.11.2010)

7.2.6 Phraseo-Lex

Keil (1997) legt mit *Phraseo-Lex* ein Lexikon vor, das die Repräsentation und die Verarbeitung verbaler Phraseologismen demonstriert. Sie beschreibt (deutsche) verbale Phraseologismen mit ihren spezifischen semantischen Eigenschaften und modelliert diese auf einem computerlinguistischen Hintergrund. *Phraseo-Lex* ist als lexikalische Datenbank konzipiert und zur Erfassung und Klassifizierung verbaler Phraseologismen geeignet. Die Tabelle (7.2) zeigt auszugsweise den Aufbau von *Phraseo-Lex* am Beispiel des Idioms «einen Streit vom Zaun brechen» (Keil, 1997, S. 222).

Die syntaktischen Merkmale der *Stabilität* beinhalten wichtige Detailinformationen für die Lexikoneinträge. Im Beispiel (Tab. 7.2) ist die *Expansion* der Komponente «Streit» durch Attribute möglich (z. B. «einen grossen Streit...»). Keil nennt diese Komponenten *semantisch autonom*. Innerhalb von Idioms gibt es austauschbare Teilkomponenten mit identifizierbaren Teilbedeutungen. Bsp. «einen (grossen, den grössten) Vogel abschiessen». Im Gegensatz dazu kann in der Redewendung «ins Gras beißen» die Komponente «Gras» nicht modifiziert werden.¹² Semantisch autonome Komponenten können z. B. durch Tests der Passivierbarkeit gefunden werden (Keil, 1997, S. 88). In einem ähnlichen Sinn wie die Expansion vermitteln weitere Stabilitätsmerkmale Informationen dazu, welche Komponenten mit Relativsätzen verbunden werden können, und welche Komponenten negiert werden können (neben der Negation des ganzen Satzes).

Die *semantischen Merkmale* beinhalten als Entsprechung zur syntaktischen Expansion Angaben zur semantischen Modifizierbarkeit, im vorliegenden Beispiel prototypische Adjektive zu «Streit». Die semantische Struktur beinhaltet thematische Rollen wie Agens und Patiens (Strukturzuweisung); diese sind durch semantische Typen (Selektionsrestriktionen) ergänzt (semantische Wertigkeit). Die Paraphrasen zu den präsentierten Phraseologismen wurden mittels *Textanalyse* gefunden, wie Keil an diversen Beispielen zeigt (S. 105–122).

Innerhalb der *pragmatischen Merkmale* fällt vor allem die *Konnotation* auf, die wichtige Aspekte der Bedeutung repräsentiert.

7.2.7 Dutch Electronic Lexicon of Multiword Expressions

Im holländischen DuELME-Projekt¹³ werden einerseits *Konstruktionsklassen* beschrieben. Diese bestehen aus Namen, PoS-tag, Mapping etc. und be-

¹²Kompositionalität von Idioms.

¹³URL: <http://duelme.inl.nl/> (28.8.2010)

Tabelle 7.2: Phraseo-Lex (nach Keil, 1997)

Merkmal	Beispiel
Lemma:	«einen Streit vom Zaun brechen»
Paraphrase:	einen Streit unbedingt beginnen
SYNTAX	
Syntaktischer Typ:	VPL1
Struktur:	[NP _{akk} [einen Streit]][PP _{vom} [vom Zaun]][V [brechen]]
Valenz:	freie Valenzen: 1; interne V: NP _{akk} , PP _{vom} ; externe V: NP _{nom}
Stabilität:	Angaben zu Passivierung (+), Imperativsatz (-), Nominalisierung (-), Negationen, Expansion durch Attribut (Streit), Quantifizierung (Streit)
Syntakt. Anomalie:	nein
Unikale Komponente:	nein
Varianten:	Keine (lexikalisch und strukturell)
SEMANTIK	
Strukturzuweisung:	jmd=Agens; einen_Streit=Patiens; vom_Zaun=keine
Klassifikation:	teilkompositionell, nicht mehrdeutig, teilmotiviert
Semantische Wertigkeit:	AGENS(+belebt, +mensch); PATIENS(+abstrakt)
Modifikation:	Streit (Adjektive)
Synonyme, Antonyme:	A: das Kriegsbeil begraben
PRAGMATIK:	
Konnotation:	wertend, kritisch
Verwendungssituation:	unspezifisch
Diaphasische Information:	normalsprachlich (Lebensphase)
Diastatische Information:	unspezifisch (schichtspezifisch, Soziolekt)
Diatopische Information:	überregional (räumlich, Dialekt)

schreiben detailliert den schematischen Aufbau der Klasse. Einzelne *Mehrwortausdrücke* werden mit Hilfe dieser Klassen beschrieben. Die Darstellung von Mehrwortausdrücken geschieht ähnlich wie im *Phrasemanager* (Pedrazzini, 1994). Weil der klassenbasierte Ansatz für die Lexikographen problematisch sein kann, wurden zusätzliche Parameter für variable Beschränkungen auf der Ausdrucksebene eingeführt (Grégoire, 2006). Die linguistische Beschreibung der Klassen und Parameter (als Patterns bezeichnet) basiert auf Konstituenten- und Abhängigkeitsstrukturen und enthält nummerierte Slots für die lexikalischen Komponenten (Komponentenliste genannt). Damit diese Patterns richtig verwendet werden können, müssen die Teile des MWE in der richtigen Reihenfolge in das Patterns eingefügt

werden. Tabelle (7.3) zeigt beispielhaft den Lexikoneintrag für «angst aanjagen» (Angst einjagen) in Klammernotation, Abbildung (7.1) zeigt die korrespondierende Baumstruktur (vgl. Luder & Clematide, 2010).

Tabelle 7.3: DuELME: Beispiel «angst aanjagen»

Expression:	angst aanjagen
Components:	EMP angst[sg] aan_jagen[part]
Pattern:	[.VP [.obj2:NP (var)] [.obj1:NP [.det:D (1)] [.hd:N (2)]] [.hd:V (3)]]
Description:	Expressions headed by a verb, taking (1) a variable indirect object and (2) a direct object consisting of a fixed determiner and an unmodifiable noun.

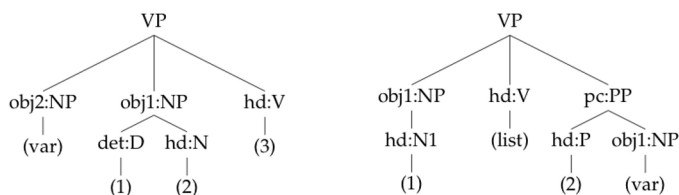


Abbildung 7.1: Baumstrukturen «angst aanjagen», «angst voor»

Komponentenliste und Pattern im Beispiel sind so zu verstehen: Angst ist ein modifizierbares direktes Objekt im Singular mit einem leeren Artikel (empty determiner (EMP)), das Verb ist ein Partikelverb (part). Die syntaktischen Kategorien sind explizit definiert. Die Aufgabe des Lexikographen besteht darin, dass richtige Pattern, die nötigen Parameter und Listen für einen neuen Eintrag zu finden und die Komponentenliste korrekt zusammenzustellen.

Der DuELME-Formalismus wurde von mit den Verbpatterns der Corpus Pattern Analysis (CPA) verglichen (Luder & Clematide, 2010), und für die aktuelle Implementierung des JAKOB-Lexikons werden die Vorteile der beiden Systeme kombiniert (Kap. 9.1.3).

7.2.8 OLIF

Hier folgt eine kurze zusammenfassende Beschreibung des OLIF-Lexikonformats, das als Grundlage für die Erarbeitung des neuen Lexikons verwen-

det wurde. Konkrete Beispiele der OLIF-Anwendung befinden sich im Kap. (9.1.2).

Das *Open Lexicon Interchange Format* (OLIF) ist ein XML-basiertes lexikalisch-terminologisches Austauschformat, das es ermöglichen soll, auf einfache Weise Daten zwischen verschiedenen Sprachtechnologie-Anwendungen auszutauschen. Der Standard wurde vom OLIF Consortium¹⁴ geschaffen, und das XML-Schema ist öffentlich verfügbar. Der Prototyp für das OLIF-Format wurde Mitte der 1990er Jahre durch verschiedene Institutionen und Firmen aus dem Bereich der Sprachtechnologie entwickelt. Die OLIF-Version v.2 ist verfügbar seit Februar 2002, eine Kurzdokumentation stammt von McCormick, Lieske und Culum (2004).¹⁵

OLIF bietet sowohl Lexikon- als auch Terminologieaspekte. OLIF-Datensätze können einerseits lemma-orientiert sein, wie dies in Wörterbüchern der Fall ist, aber auch konzeptorientiert, wie dies in Terminologiedatenbanken der Fall ist. OLIF ist «wortsinnorientiert» (McCormick et al., 2004), d. h. ein Eintrag gleicht einer *Lexical Unit* sensu FrameNet und beherbergt eine Kollektion von Merkmalen zu einem spezifischen Wort- oder Phrasensinn, mit optionalen Verknüpfungen (McCormick, 2005). Das OLIF-Schema bietet eine grosse Zahl an Eigenschaften (Attributen) zur Spezifikation der Einträge an, sowohl zu morphologischen, syntaktischen und semantischen als auch zu pragmatischen und administrativen Kategorien; ausserdem gibt es ausführliche Verlinkungsmöglichkeiten zwischen den Einträgen, z. B. für die Übersetzung in andere Sprachen und für Querverweise (cross references). Die Datenkategorien wurden im Anschluss an bestehende Sprachtechnologiesysteme (z. B. Wörterbücher, maschinelle Übersetzungssysteme) sowie mit Rücksicht auf den Standard ISO 12620 bestimmt. Ziel der Kategoriendefinition in OLIF war es, das Kategoriensystem auf möglichst viele bestehende *MT-Systeme* (machine translation) abbilden zu können.

Die möglichen Attributwerte sind für gewisse Kategorien mehr oder weniger im OLIF-Schema vorgegeben, wie z. B. für die syntaktischen Eigenschaften Kasus, Numerus und Person; für andere Kategorien ist der Katalog der möglichen Werte offener und nicht eindeutig, wie z. B. für die Kategorien <subject field> oder <semantic reading>.

Lexikoneinträge können mit verschiedenen übergeordneten Strukturen verknüpft werden: Die <concept-ID> verknüpft einzelne OLIF-Einträge mit einem übergeordneten Konzept; so ist es möglich, mehrere Einträge mit gleicher Bedeutung dem gleichen Konzept zuzuordnen (vgl. Synset in WordNet

¹⁴URL: <http://www.olif.net/consortium.htm> (15.11.2010)

¹⁵URL: <http://www.olif.net/> (15.11.2010)

oder Semantic Frames in FrameNet).

Das OLIF-Schema schien für eine Neugestaltung des JAKOB-Lexikons geeignet und entsprach auch computerlinguistischen Anforderungen. Da für die OLIF-Kategorien in den im vorliegenden Projekt relevanten Bereichen (Konstruktionsgrammatik, Psychotherapie, Gesprochene Sprache) keine praktischen Beispiele vorhanden sind, und weil die OLIF-Kategorien oft nur rudimentär beschrieben und unvollständig sind, haben wir uns entschlossen, das OLIF-Schema bei Bedarf mit eigenen Werten zu ergänzen. Dadurch ist allerdings die aktuelle Lexikonimplementierung nicht mehr in allen Einzelheiten mit dem OLIF-Schema kompatibel.

7.2.9 Valenzwörterbuch deutscher Verben: VALBU

VALBU ist ein einsprachiges Wörterbuch deutscher Verben (Schumacher et al., 2004). Ziel dieses Projekts war die Erarbeitung eines didaktisch orientierten *Valenzwörterbuchs deutscher Verben*. Das Wörterbuch enthält eine umfassende semantische und syntaktische Beschreibung von 638 Verben mit ihrer spezifischen Umgebung, ferner Informationen zur Morphologie, Wortbildung, Passivfähigkeit, Phraseologie und Stilistik sowie zahlreiche Verwendungsbeispiele. Das VALBU ist auch als E-VALBU online verfügbar.¹⁶

Methodik, Datengrundlage: Die Stichwortauswahl von 638 Verben lehnt sich an die Wortschatzliste des *Zertifikats Deutsch* an und richtet sich nach den Erfordernissen des Unterrichts in Deutsch als Fremdsprache. Die heuristischen Verfahren sind durch die Textkorpora des IDS motiviert. Die Verwendungsbeispiele sind zum größten Teil diesen Textkorpora entnommen, teilweise aus didaktischen Gründen in adaptierter Form.

Semantische Kategorien: Als weitere Möglichkeit, die eng mit der vorher erwähnten zusammenhängt, wird die Klassifikation der syntaktischen Komplemente in semantische Kategorien aufgeführt, wie sie im VALBU eingeführt wird. Die Kategorien entsprechen in etwa den semantischen Typen.

Ebenfalls im VALBU gibt es zusätzlich eine verbspezifische Beschreibung der beteiligten «Teilnehmer» (verb-specific description of participants). Auch dies ist eine Ergänzung zum semantischen Typ und beschreibt weitere Eigenschaften auf ein bestimmtes Verb bezogen.

¹⁶URL: <http://hypermedia2.ids-mannheim.de/evalbu/index.html>

7.3 Ontologien, Texttypen, Diskursmuster

Durch die Erweiterung der *Konstruktionsgrammatik* zu einer *Diskursgrammatik* reichen die semantischen Eigenschaften und Restriktionen, wie sie auf der Satzebene wichtig sind, nicht mehr aus. Mit der Erweiterung zum Diskurs geraten Diskurseigenschaften, die pragmatischer Natur sind, also extralinguistische Merkmale betreffen, in den Fokus der Aufmerksamkeit. Der Ontologiebegriff ist deshalb nicht nur wort- und satzsemantisch gemeint, sondern bezieht sich ebenso auf die Ebene der Situations- und Kommunikationsbedingungen im Diskurs.

Im Folgenden sollen semantische und pragmatische Typologien und Ontologien geprüft werden, die einen praktischen Nutzen für das Lexikon leisten könnten, d. h. die dazu beitragen, Lexikoneinträge möglichst eindeutig zu charakterisieren. Ausserdem muss die Frage geklärt werden, welche Eigenschaften überhaupt dem Lexikoneintrag zugeordnet werden können; Eigenschaften wie «Gesprächssorte» beziehen sich auf ganze Texte/Gespräche, oder zumindest auf grössere Gesprächsausschnitte. Das entsprechende Merkmal im Lexikoneintrag würde sich dann darauf beziehen, dass die entsprechende Konstruktion nur oder vorzugsweise in einer bestimmten Gesprächssorte vorkommt oder dass sie in dieser Textsorte eine spezifische Bedeutung hat. Dazu stellt sich die Frage, auf welcher Ebene eine Typologisierung stattfinden sollte. Sager (2001) unterscheidet die Makro-, Meso- und Mikroebene, wie in der Tabelle (7.4) dargestellt (nach Sager, 2001). Er unterscheidet eine eher abstrakte theoretische Dreiteilung gegenüber einer empirisch-konkreten Begrifflichkeit:

Tabelle 7.4: Gesprächstypologie nach Sager (2001)

Theoretisch logisch abstrakt	Empirisch praktisch konkret
Gesprächstyp	Gesprächssorte
Gesprächsmuster (Gesprächsphase)	Kommunikative Gattung (Episode, thematischer Abschnitt)
Gesprächsakt	Kommunikative Funktion

Ziel der folgenden Beschreibungen ist nicht eine möglichst vollständige Typologisierung, sondern eine Entscheidungsgrundlage für die Definition und Formulierung der vorgesehenen neuen pragmatischen Merkmale für das Lexikonprojekt.

7.3.1 Makroebene: Domain, Genre, Subject Field

Der Begriff *Domain* wird im Folgenden für die Makroebene der Gesprächstypologie (Sager, 2001) eingesetzt, nämlich für den *Gesprächstyp* oder die *Gesprächssorte*. Diese Begriffe sind Bezeichnungen für ganze Gespräche/Texte und entsprechen daher am ehesten dem Merkmal <subjField>, wie es im OLIF-Schema verwendet wird. In OLIF sind damit klar abgegrenzte, durch den alltäglichen praktischen Sprachgebrauch (vor allem in schriftlichen Texten) bestimmte Sachgebiete bezeichnet (McCormick, 2005). Beispiele für die thematische Ausrichtung und die Abstraktionsebene dieser Begriffe sind in Tabelle (7.5) beschrieben. OLIF bietet insgesamt 38 Eigenschaften in dieser Liste an.¹⁷

Tabelle 7.5: OLIF-Eigenschaft <subjField>

Value	Description
agriculture	farming and agriculture
aviation	aviation and aerospace
botany/zoology	botany and zoology
fisheries	fishery science and technology
medicine	medicine
social	social science and policy etc.

Der Begriff *Genre* bezeichnet oft dasselbe wie *Domain* (Lee, 2001).¹⁸ Diese Eigenschaft basiert auf externen, nicht-linguistischen Kriterien: wo- zu findet das Gespräch statt und wozu wird es gebraucht? Genre ist eine Eigenschaft für ganze Texte oder ganze Gespräche:

a category assigned on the basis of external criteria such as intended audience, purpose, and activity type, that is, it refers to a conventional, culturally recognised grouping [of texts]. (Lee, 2001, S. 38)

Ferguson (1994, S. 21) bezeichnet Genre als einen *Mitteilungstyp*, der konventionalisiert innerhalb einer community verwendet wird. Als Beispiele führt er an: Rätsel, Psalm, Gesetzesdokument, Chat, Debatte, Liebesbriefe,

¹⁷Die meisten dieser Items sind für die vorliegenden Daten – Gesprächstranskripte – nicht sehr relevant; die Auswahl könnte allenfalls fest auf «Therapiegespräch» gesetzt werden.

¹⁸(Lee, 2001) beschreibt die Kategorien Genre, Register, Text type, Domain und Style in seiner Übersicht «Clarifying the concepts and navigating a path through the BNC jungle».

Gebete, Rezepte, Todesanzeigen etc. Auch Östman (2005) erwähnt ähnliche Beispiele, wie Rezepte, Tischgespräche, Märchen, Arztgespräch und ordnet dem Genre die Rolle des «contextual setting» zu.

Kuiper (2009, S. 16) versteht unter *formulaic genres* beispielsweise Verlobungsanzeigen und Eishockeyspiel-Kommentare. Er liefert keine klare Definition, betont aber die kommunikative Funktion, die durch das Genre bestimmt wird. Bergmann und Luckmann (1995) bezeichnen *kommunikative Genres* als Lösungen für spezifische kommunikative Aufgaben («communicative genres are solutions to specifically communicative problems»).

7.3.2 Meso- und Mikroebene: Gesprächstypen und lokale Muster

Text Type

Der *Text Type* bezeichnet gemäss Lee (2001) eigentlich *interne linguistische Charakteristiken* eines Gesprächsabschnittes; er erweitert diesen Begriff aber etwas zu Kategorien wie *Narrativ*, *Beschreibung* (description), *Exposition*, *Argumentation*. Im *Metzler Lexikon Sprache* gibt es zum Texttyp unterschiedliche Definitionen; eine davon deckt sich mit der von Lee erwähnten: Texttypen sind *Muster sprachlichen Handelns* (Glück, 2005). Auch Östman unterscheidet argumentative, narrative, instruktive, expositorische und deskriptive Texttypen. Die Unterscheidung gilt sowohl für Text (geschrieben) als auch für den gesprochenen Diskurs. Bezüglich der Situationsabhängigkeit linguistischer Phänomene sprechen Deppermann und Spranz-Fogasy (2001) von *Modalitäten* des Gesprächs (z. B. Spass, Nacherzählung).¹⁹

In der Erzählanalyse JAKOB, die sich grundsätzlich mit Alltagserzählungen befasst, werden formal «prototypische», d. h. *episodische Erzählungen*, iterative Erzählungen (Erzählung eines immer wieder ablaufenden Musters) und Karriereerzählungen (Bericht über eine langfristige Entwicklung) unterschieden (Arboleda et al., 2010).

Diskurspattern

Wie Form und Bedeutung in der Konstruktion zusammengeführt werden und eine Einheit bilden, werden gemäss Östman (2005, S. 130) analog da-

¹⁹Für die Erzählanalyse JAKOB besteht die Konvention, dass der Begriff *Modalität* für die Beschreibung von *Textsegmenten* reserviert ist (Segmente sind Sätze, einfache Subjekt-Prädikatverbindungen). *Modalität* bezeichnet in diesem Zusammenhang Satzart und Bedeutungsmodi für das Segments: Aussagesatz, Fragesatz, Ausrufesatz, Negation, Passiv, Modalverben.

zu Genre und Text-Typ im *Diskurspattern* repräsentiert. Diskurs-Patterns sind in dieser Sichtweise konventionalisierte Assoziationen zwischen Text-Typ und Genre. Östman verwendet <discourse pattern> als Merkmal, um Konstruktionen zu beschreiben. Östmans Vorschläge für die Darstellung des Diskurspatterns als Merkmal der Konstruktion, das über die Grenze der Satzkonstruktion hinausgeht (dp) sind in der Abbildung (7.2) dargestellt. Alternativen sind die Darstellung in der Konstruktionsbox selbst (a), darum herum (b) oder in einer separaten Box (c) (vgl. CxG-Formalismen).

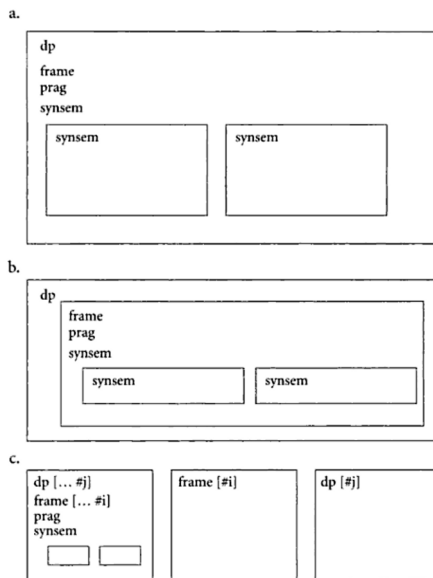


Abbildung 7.2: Diskurspattern (dp) und Konstruktion (Östman, 2005)

Register, Stil, Sublanguage

Register charakterisieren gemäss Lee (2001, S. 41) die Sprachverwendung in *sozialen Situationen*. Register sind z. B. durch Eigenschaften wie «formell», «informell», «religiös», «umgangssprachlich», «humoristisch» etc. gekennzeichnet. Sie sind eine Eigenschaft lokaler Patterns und werden abhängig von Situation und Adressaten eingesetzt. Andere Autoren verwenden für diese Eigenschaften die Bezeichnungen *Kode* oder *Stil* (z. B. Deppermann

& Spranz-Fogasy, 2001), die abhängig vom lokalen Gesprächsthema eingesetzt werden (z. B. formell vs. informell).

Biber und Finegan (1994a, S. 4) stimmen der Definition von Lee zu («language variety viewed with respect to its context of use (social situation)»). Für den situationsabhängigen Gebrauch (situational uses) postulieren sie aber ebenso die Begriffe *genre*, *text type* und *style*.

Im Buch «Sociolinguistic perspectives on register» beschreiben Biber und Finegan (1994a) das linguistische Repertoire von Sprachgemeinschaften als Inventar der verwendeten *Register*. Sie unterscheiden Variationen im Register gegenüber Variationen im *Soziolekt* (social dialect variation). Sie setzen die Begriffe register, style, genre, text type und dialect sowie social/situational aspect in Beziehung zu einander, beklagen aber gleichzeitig den Mangel an ausgearbeiteten Typologien («Lack of a paradigmatic, or even dominant, methodological and theoretical framework for the analysis of registers!» S. 6).

Ferguson (1994) verwendet die Begriffe *dialect*, *register*, *genre* und *conversation*. Wesentliche Aspekte der menschlichen Sprache sind durch Konvention verursacht, nämlich durch die Art und Weise, mit der sich die Mitglieder einer sozialen Gemeinschaft gegenseitig die Kombination von lautlichem Sprachausdruck und Bedeutung kenntlich machen (Ferguson, 1994, S. 15). Kommunikationssituationen, die regelmässig in einer Gesellschaft vorkommen, bewirken Sprachmerkmale und -strukturen, die z. B. die schnelle Kommunikation optimieren oder die Zugehörigkeit zu einer Gemeinschaft bestärken (S. 20). Sprachbenutzer können in der verbalen Interaktion die kommunikativen Ressourcen ausschöpfen und Variationen kreativ einsetzen. Eine der unbeantworteten Fragen ist gemäss Ferguson (1994, S. 27), wie die *Konventionalisierung* zustande kommt, wie Gruppen von Sprachbenutzern dazu kommen, gewisse Eigenschaften von Sprachgebrauch und -struktur zu teilen. Die Theorie des *lexical priming* von (Hoey, 2005) könnte darauf eine Antwort geben (Kap. 6.12).

Das Merkmal *style* wird oft mit einer ähnlichen Bedeutung versehen wie das oben beschriebene *register*. Lee (2001, S. 45) zählt dazu einerseits Gegensatzpaare wie «formal – informal», «one-way – interactive», «considered – impromptu», und andererseits Kategorien wie «colloquial» (umgangssprachlich), «intimate», «humorous» etc.

Die Kategorie *Sublanguage* stammt aus der Computerlinguistik und beinhaltet das Spezialvokabular ganz bestimmter Wissensdomänen, wie es z. B. für sprachverstehende Systeme angewendet wird (air traffic control, weather report etc.). Sublanguage ist der Kategorie Genre/Domain ähnlich, aber enger eingeschränkt auf ein Spezialgebiet (S. 38).

Topic

Gemäss *Metzler Lexikon Sprache* (Glück, 2005) ist Topik das lokale Thema, die «pragmatische Rolle des Gegenstandes, von dem aus ein Geschehen perspektiviert wird» (funktionale Satzperspektive). Lee (2001) verortet das Merkmal *topic* in verschiedenen Rollen und Funktionen, eine Tendenz besteht aber darin, *topic* als *Thema* aufzufassen, als Beschreibung des *Inhalts* einer sprachlichen Einheit. So präsentiert er eine Studie, die *topic* als eines der Attribute von Genre sieht:

[Genre characterised as] a set of seven attributes: domain (e.g., art, science, religion, government), medium (e.g., spoken, written, electronic), content (topics, themes), form (e.g., generic superstructures [...] or other text-structural patterns), function (e.g., informative, persuasive, instructive), type (the rhetorical categories of narrative, argumentation, description, and exposition) and language (linguistic characteristics: register/style [...]). (Steen, 1999, zit. nach Lee, 2001, S. 49)

In erzähl- und gesprächsanalytischen Untersuchungen wird der Begriff *topic* verwendet, um das aktuelle Gesprächsthema (z. B. in einer Therapie-sitzung) zu benennen, über das die Beteiligten in einem kleinen Gesprächsausschnitt (deshalb auch lokales Thema) gerade sprechen. Um die Gesprächsorganisation und den Ablauf zu beschreiben, werden narrative Techniken (narrative devices) verwendet, die mit Begriffen wie *topic proposal*, *topic ratification*, *topic shift* oder *topicalization* bezeichnet werden (Kern & Quasthoff, 2005, S. 20f).

Das Merkmal *topic* wird für die Neugestaltung des Lexikons verwendet und beschreibt thematische Eigenschaften und Restriktionen eines Lexikon-eintrags. Dieses Merkmal ist allerdings bis jetzt nur sehr vage formuliert und enthält heterogene Werte wie <streit>, <familie>, <gefühle> etc., die neben den semantischen Eigenschaften und den pragmatischen Merkmalen der Situation und des Kommunikationskontextes inhaltliche Hinweise zum Gebrauch eines Ausdrucks liefern sollten (vgl. 9.1.3).

Weitere Gesprächsmerkmale

Die *Kommunikative Gattung* (Sager, 2001) kann ebenfalls als eine Bezeichnung für Gesprächstypen angesehen werden. Kommunikative Gattung, Gesprächsmuster und Gesprächsphasen bezeichnen Einheiten unterhalb der Gesprächsebene. Kommunikative Muster und Gattungen gelten als zentra-

le Einheiten der Sprache und Kommunikation; die Sicht auf diese Phänomene erfolgt aber hier eher aus der Perspektive von Gesprächsanalyse und Gesprochener Sprache, im folgenden Zitat in Verbindung gebracht zu den Theorien der Konstruktionsgrammatik (Günthner, 2006b, S. 16):

Sprachliches Handeln geschieht nicht allein auf der Grundlage abstrakter Regeln, die stets neu kombiniert werden, sondern zu einem erheblichen Teil durch die Reaktivierung memorisierter, im Wissensvorrat gespeicherter, konventionalisierter Konstruktionen und Muster.

Konstruktionen und kommunikative Gattungen stellen holistische Gestalten dar, die nicht als autonome Module sondern im Zusammentreffen von prosodischen, morphosyntaktischen, semantischen, stilistischen, interaktional-sequenziellen und kontextuellen Faktoren zu analysieren sind. Aufgrund ihrer gestalthaften Qualität erlauben sie Projektionen bzgl. des weiteren Verlaufs.

Vorgeformte Muster erleichtern – wie jede Form der Routinisierung oder Institutionalisierung – die Kommunikation.

Sie sind Teil des sprachlichen und soziokulturellen Wissensvorrats.

Sie werden in aktuellen Kommunikationssituationen re-aktiviert und sind folglich dem dialogischen und prozessualen Charakter mündlicher Interaktion unterworfen.

Das *SPEAKING-Modell* von Dell Hymes beschreibt *Sprechakte* detailliert mit 16 Komponenten der linguistischen Interaktion, die theoretisch für alle Diskurstypen gültig sind (Deppermann & Spranz-Fogasy, 2001; Sager, 2001). Später subsumierte er die 16 Komponenten unter acht Gruppen, die das Akronym «SPEAKING» ergeben, wie im Folgenden gezeigt (Hymes, 1972, nach Sager, 2001, S. 1468):

S – setting and scene : physikalischer Ort und Zeitpunkt des Sprechaktes, psychologische und sozio-kulturelle Interpretation des setting.

P – participants : Sprecher, (Ab)Sender, Hörer, Rezipient, Auditorium, Adressat.

E – ends : Zweck, Ergebnis, Ziel, Absicht.

A – act sequence : Art der Äusserung, Inhalt der Äusserung.

K – key : Ton, Modus, Charakter der Äusserung.

I – instrumentalities : Kanal, Kode, Formen des Sprechens.

N – norms : Normen/Konventionen der Interaktion, Normen/Konventionen der Interpretation.

G – genres : Gattungen der Kommunikation.

Das SPEAKING-Modell fokussiert auf die breit angelegte Darstellung dessen, was bei der Produktion eines Sprechaktes abläuft und ist deswegen für die Beschreibung von Lexikoneinträgen viel zu ausführlich.

Die in diesem Kapitel beschriebenen Typologien und Charakterisierungen sind alles andere als einheitlich, über die Definitionen der einzelnen Begriffe gibt es keinen Konsens. Aus den vorliegenden Möglichkeiten werden für das JAKOB-Lexikon vier Ebenen der Beschreibung übernommen: *Domain (subject field)*, *Texttyp*, *Register* und *Topic*. Es wird ein Versuch unternommen, diese vier Kategorien der pragmatischen Beschreibung für das vorliegende Projekt möglichst prägnant zu beschreiben (siehe Kap. 9.1.3). Allerdings wird sich zeigen, dass die Beschreibung und vor allem die Anwendung dieser Kategorien dem Lexikographen grosse Schwierigkeiten bereitet.

7.3.3 Semantische Ontologien

Ontologien dienen der Repräsentation von konzeptuellem Wissen, wobei zwischen allgemeinen und domänenspezifischen Ontologien unterschieden werden kann. *WordNet* (siehe Kap. 7.2.1) kann mit seinen Hyperonym- und Hyponym-Relationen, die die Synsets verbinden, als allgemeine Ontologie bezeichnet werden. Da jedoch WordNet nicht systematisch als Ontologie entwickelt wurde, existieren viele semantische Inkonsistenzen und Redundanzen. Wie aus dem nachstehenden Beispiel (38) unschwer zu erkennen ist, geht ausserdem die Detaillierung der Kategorien weit über das hinaus, was eine praktisch anwendbare Ontologie für die Typisierung der Argumente der Verbpattern leisten muss (vgl. auch die Kritik an den WordNet-Kategorien von Hanks & Pustejovsky, 2005).

- (38) Hierarchisch angeordnete Synsets für das Nomen «dog» (IS_A-Relationen):

```

dog, domestic dog, Canis familiaris
=> canine, canid
=> carnivore
=> placental, placental mammal, eutherian, eutherian mammal
=> mammal
=> vertebrate, craniate
=> chordate
=> animal, animate being, beast, brute, creature, fauna
=> ... (Quelle: Wikipedia)

```

CPA-Ontologie

Jezek und Hanks (2010, S. 9) beschreiben verschiedene Bedeutungen des Begriffs *Ontologie*; im Kontext der Philosophie beschäftigt sich Ontologie mit den Ordnungs- und Begriffsbestimmungen des Seins. Im Kontext der *Wissenrepräsentation*, um den es hier geht, wird unter Ontologie eine «explizite Spezifikation einer gemeinsam Konzeptualisierung» verstanden, d. h. eine sprachunabhängige Repräsentation der Welt oder eines Ausschnitts der Welt, die von den spezifischen Handlungen eines Agenten und spezifischen Situationen abstrahiert (Carstensen et al., 2004, S. 451). Im Kontext der *Corpus Pattern Analysis* (CPA) formulieren Jezek und Hanks (2010) die Aufgabe der Ontologie so:

...an ordered collection of the contentful terms of a language, both abstract and concrete. This is the sense in which the term ontology is applied for instance to WordNet. It is the basis of the sense in which we shall use the term in this paper, but, rather than engaging in armchair speculation about semantic relations and synonym set, we use corpus evidence to try to find out which words are collocates of each other and how different collocations are associated with different meanings. (S. 9)

Die CPA-Ontologie ist aus der *Brandeis Ontology* entstanden (Pustejovsky et al., 2004)²⁰ und basiert auf Pustejovskys Arbeiten zum *Generativen Lexikon* (Pustejovsky, 1995).

²⁰ «The Brandeis Shallow Ontology (BSO) is a shallow hierarchy of types selected for their prevalence in manually identified selection context patterns. At the time of writing, there are just 65 types, in terms of which patterns for the first one hundred verbs have been analyzed. New types are added occasionally, but only when all possibilities of using existing types prove inadequate. Once the set of manually extracted patterns is sufficient, the type system will be re-populated and become pattern-driven» (Pustejovsky et al., 2004, S. 4).

Eine Ontologie muss genug Differenzierungsmöglichkeiten bieten, so dass für alle Argumentpositionen in den Verbpatterns aussagekräftige (und distinktive) semantische Typen zur Verfügung stehen. Die Ontologie, die hinter den semantischen Typen von OLIF steht, ist für die Umsetzung der Verbpatterns im Lexikon nicht passend, deshalb wird für die Beschreibung der Verbpatterns die CPA-Ontologie verwendet, wie sie aktuell auch im *Pattern Dictionary of English Verbs* (PDEV) verwendet wird.²¹

Ein kleiner Auszug aus der CPA-Ontologie ist in Tabelle (7.6) abgebildet. Die oberste Ebene besteht aus dem Begriff Anything, der sämtlichen Kategorien übergeordnet ist. Die praktische Anwendung der CPA-Ontologie im Lexikon ist im Kap. (9.1.3) beschrieben.

Tabelle 7.6: Auszug aus der CPA-Ontologie (Ebene 1: Anything)

Ebene 2	Ebene 3	Ebene 4	Ebene 5	Ebene 6
Entity	Physical Object	Inanimate Animate	Artwork Human Animal	
Eventuality	State Event	Abstract Activity	Psych Action Speech Act	Attitude Emotion Goal Request Question

7.4 Diverse Konzepte zum Lexikon

7.4.1 Nominalität vs. Verbalität

In seinem Buch «Lexikon und Grammatik in Interaktion» gibt Weber (2010) einen historischen und theoretischen Überblick über *lexikalische Kategorisierungsprozesse*. Er erwägt verschiedene Kategorisierungsmöglichkeiten als konzeptuelle Alternativen zu den syntaktischen Begrifflichkeiten von *Wortarten* und *Redeteilen*. Die lexikalischen Kategorien des *Sprechens* sind abzugrenzen von den kognitiv-semantischen Kategorien des *Denkens* einerseits und von den ontischen Kategorien des *Seins*. Dies entspricht den Grundannahmen *kognitiver Ansätze* einer Theorie lexikalischer Kategorisierung (Weber, 2010, S. 45). Die *diskursfunktionale Konzeption* lexikalischer Kategorisierung verbindet dagegen Lexikon und Grammatik. Sprecher pro-

²¹ URL: <http://deb.fi.muni.cz/pdev> (15.11.2010)

duzieren und verstehen nicht Sätze, die syntaktischen etc. Regeln genügen, sondern kontextualisierte Äusserungen! Lexikalische Einheiten sind Syntagmen, die als Basis für Äusserungen zur Verfügung stehen (S. 84f).

Gemäss Hopper und Thompson (1984) widerspiegeln die lexikalischen Kategorien typische *Diskursfunktionen*: Prototypische Diskursfunktionen bestehen (1) aus einem «discourse-manipulable participant» und (2) einem «reported event». (1) sind *Substantive*, d. h. Objekte. Substantive werden referenziell gebraucht; sie werden einmal erwähnt, dann erfolgt ein anaphorischer Bezug darauf. Substantive führen in der Regel *neue Referenten* ein, d. h. neue, noch nicht aktive Information. Bereits aktive Information wird hingegen durch *Pronomen* dargestellt. Die daraus resultierende Satzstruktur wird von Du Bois, Kumpf und Ashby (2003) als *preferred argument structure* (PAS) bezeichnet, die sich durch folgende Beschränkungen auszeichnet:

1. *One Lexical argument constraint*: Äusserungen (Sätze) in der gesprochenen Sprache weisen selten mehr als eine Nominalphrase mit substantivischem (d. h. lexikalisiertem) Kern auf (Avoid more than one lexical core argument).
2. *Non Lexical A Constraint*: Lexikalisierte Nominalphrasen sind nicht über alle Argumentstellen verteilt, sondern werden präferiert als Objekt transitiver Sätze und als Subjekt intransitiver Sätze verwendet und kaum als Subjekt transitiver oder ditransitiver Sätze.
3. *One new argument constraint* bezieht sich auf die funktionale Perspektive: Ein Satz weist selten mehr als ein Argument auf, das sich auf einen neuen (noch nicht aktiven) Gegenstand bezieht.
4. *Given A Constraint*: Subjekte intransitiver Sätze (A) sind überwiegend pronominal, oder in gewissen Sprachen durch Affixe realisiert (Du Bois, 2003, zit. nach Weber, 2010, S. 88)

Dies führt Weber zu interessanten, das Lexikon und die Lexikalisierung betreffenden Überlegungen: *Substantive* (nominal markierte Wortformen) kommen vor allem als *Erstreferenz* vor und führen neue Referenten oder Themen ein. *Verben* (verbal markierte Wortformen) prädisieren Ereignisse oder Handlungen. Am Beispiel des Wortes «Kanzler» zeigt Weber, dass diverse Ausprägungen des Wortstamms in unterschiedlichen Kontexten existieren, von klar *substantivischem Referenten* («der Kanzler») bis zum Verb «kanzeln» am andern Ende der Skala (Weber, 2010, S. 93). Weber be-

zeichnet *Nominalität* und *Verbalität* als Extrempunkte auf einem gedachten Kontinuum (siehe Abbildung 7.3).

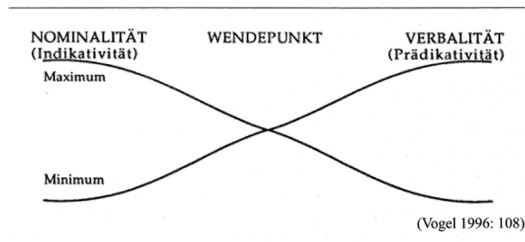


Abbildung 7.3: Das Kontinuum von Nominalität und Verbalität

Der kategoriale Status zwischen Nomen und Verb und das Verwendungspotenzial im mentalen Lexikon verändern sich permanent in einem adaptiven Prozess. Die Begriffe Nominalität vs. Verbalität widerspiegeln sich analog im Begriffspaar *Thematizität* vs. *Prädikativität* (Weber, 2010, S. 102–105). Für die Projektumsetzung und die Erfassung der Konstruktionen im elektronischen Lexikon bleiben wir aber beim Wort als kleinste Einheit und setzen nicht Wortstämme als Einheiten des Lexikons ein.

7.4.2 Konnotation im Lexikon

Sollen in einem Lexikon für automatische Sprachverarbeitung auch Bewertungen eingebaut werden? Braasch und Pedersen (2010) stellen entsprechende Ergänzungen aus dem dänischen WordNet vor. Die Autorinnen bezeichnen *Konnotation* als das Set möglicher assoziativer Bedeutungsvarianten, die ein Lexem neben seiner primären Bedeutung hat:

Connotation is seen as the set of associations implied by a lexeme in addition to its primary, literary meaning; it is evoked by one (or more) particular feature of the entity referred to and suggests attitudes, emotions and opinions like admiration or disapproval. (S. 282)

Konnotation ist eine zusätzliche Dimension zur Kernbedeutung eines Lexems (von den Autorinnen als literal meaning bezeichnet); sie drückt die Haltung des Sprechers gegenüber dem Konzept aus, das hinter dem Lexem steht. Das kann eine Emotion sein, Bewunderung, Verachtung, Geringschätzung etc. (S. 284). Die Konnotation spielt daher eine wichtige Rolle bei der

Differenzierung von Synsets. Ein Lemma hat verschiedene Subjektivitätswerte abhängig vom kommunikativen Kontext; und nur Wortbedeutungen mit gleicher Polarität (der Konnotation) sollten im gleichen Synset vorkommen (39).

- (39) a. S1: ‹Alkoholiker› ohne Konnotation: medizinischer oder psychopathologischer Begriff
 b. S2: ‹Alkoholiker›, Konnotation mit negativer Polarität: Trunkenbold, Säufer, ...
 c. S3: ‹Alkoholiker›, Konnotation mit negativer Polarität: Penner, Alki, ...

Der ‹Alkoholiker› ohne Konnotation (objective, non opinionated) ist ein Hyperonym zu einem oder mehreren (hier meistens negativen) Synsets: (1. Säufer, Trunkenbold,...), (2. Penner, Alki,...); diese beiden Synsets aus dem umgangssprachlichen Register unterscheiden sich aber in der Bedeutung. Während Synset 1) sich auf die Tatsache des Alkoholmissbrauchs bezieht, hat Synset 2) eine zusätzliche Konnotation zu ‹Obdachloser› (Braasch & Pedersen, 2010, S. 285).²²

Es ist eine kleine Anzahl von prototypischen Eigenschaften (features), die positive oder negative Konnotationen bewirken und für die entsprechende Markierung der Synsets verwendet werden. Braasch und Pedersen (2010, S. 289) nennen die folgende Liste, wobei es interessant ist, dass diese Eigenschaften bei Männern und Frauen anders bewertet werden, sowohl was ihre relative Wichtigkeit betrifft als teilweise auch in Bezug auf die Polarität der Konnotation.

- Appearance/shape/stature
- Age/experience/maturity
- Temper/mind/character
- Status/function/efforts
- Physical power
- Intellect/ability

Die Autorinnen demonstrieren die Nützlichkeit von konnotativer Information im Lexikon, die über die pure Denotation hinausgeht. Konnotation hilft beim Differenzieren von Synsets, die Nuancen des gleichen Begriff bezeichnen.

²²Die Beispiele stammen aus dem dänischen WordNet und beruhen auf Korpusevidenzen.

7.5 Fazit zum Kapitel Lexikographie

Aus der Sichtung der verschiedenen existierenden Mehrwortlexika werden folgende Anregungen für das Lexikonprojekt übernommen:

- *WordNet* zeigt mit den *Synsets* eine nützliche Kategorisierungsmöglichkeit für die Nomen auf. Allerdings verwenden wir für die Kategorisierung die CPA-Ontologie. Nomen mit demselben semantischen Typ bilden ein *lexical set* (siehe Kapitel 7.3.3).
- Die *Usuellen Wortverbindungen* (UWV-Analysemodell) liefern Informationen zu häufig vorkommenden Gebrauchsmustern von Wortverbindungen und organisieren diese in Wortbindungsfeldern um eine Kernvariante. Lexikalische Realisierungen in den Korpus-texten werden zu Wortverbindungen zusammengefasst und weiter zu Wortbindungsmustern abstrahiert. Konkrete Äusserungen werden auf dem Hintergrund dieser Muster verstanden. Das UVW-Modell zeigt interessante Aspekte des Sprachverstehens auf, hat aber keine direkten Auswirkungen auf die Umsetzung (Steyer & Brunner, 2009).
- Die *Idiomatisierten Verbalen Phraseologismen* (IVP, Söhn, 2006) zeichnen sich durch einige interessante Lexikonmerkmale aus, wie Dekomponierbarkeit, unikale Elemente und Restriktionen, wie Passivierbarkeit und Art der Negation. Diese Merkmale werden jedoch nicht weiter verfolgt.
- *Phraseo-Lex* (Keil, 1997) beinhaltet eine Merkmalsstruktur, die grosse Ähnlichkeit mit dem erweiterten JAKOB-Lexikon hat, wie zum Beispiel die Aufnahme von semantischen und *pragmatischen Merkmalen*. Innerhalb der pragmatischen Merkmale fällt vor allem die *Konnotation* auf, die besondere Aspekte der Bedeutung repräsentiert, wie sie oben auch für das dänische WordNet beschrieben sind.
- Das *DuELME-Projekt* und die darin verwendeten Formalismen dienen als Vorbild für das neue Lexikonmerkmal *bauplan*, das im Kapitel (9.1.3) beschrieben wird und das für das JAKOB-Lexikon eine wichtige Funktion übernimmt.
- Die *OLIF*-Struktur wird als Vorlage für das neue Lexikon übernommen (Kapitel 9.1.2).
- *VALBU* dient zu Konsultationszwecken; als sense-enumeratives Lexikon entspricht es aber nicht der vorgesehenen Implementierung.

Die im Kapitel zu *Genres, Texttypen und Diskursmustern* beschriebenen Begriffe und Typologien sind sehr heterogen. Aus den vorliegenden Möglichkeiten werden schliesslich für das JAKOB-Lexikon vier Ebenen der Beschreibung übernommen: *Domain (subject field)*, *Texttyp*, *Register* und *Topic*. Eine ausführliche Beschreibung und Charakterisierung der implementierten Merkmale erfolgt im Kapitel (9.1.3).

Für die semantische Typisierung der lexikalischen Items wird die im Kapitel (7.3.3) beschriebene CPA-Ontologie übernommen. Sie erlaubt die semantische Kategorisierung der Nomen und ermöglicht es damit, die Verben und ihre Gebrauchsmuster (Verbpattern) im Lexikon mit aussagekräftigen semantischen Informationen zu versehen. Zusätzlich wird für die Verben die semantische Kategorisierung mit den OLIF-Werten beibehalten.

8 Daten und Methoden

Das folgende Kapitel stellt die *Gesprächs- und Korpusdaten* vor, die als Grundlagen für die Lexikonergänzungen im vorliegenden Projekt dienen. Es folgt eine Auflistung von bestehenden Lexika und MWE-Sammlungen, die für die lexikographische Arbeit als Referenz und zu Vergleichszwecken verwendet werden und die z.T. schon im vorhergehenden Kapitel beschrieben wurden. Ausserdem werden die *Methoden* vorgestellt, mit denen die Gesprächstranskripte und die verwendeten Korpora durchsucht und analysiert wurden, um daraus die Lexikoneinträge abzuleiten. Im Wesentlichen sind dies *gesprächsanalytische* und *korpuslinguistische* Verfahren.

Die vorliegende Arbeit stützt sich in erster Linie auf Daten aus Psychotherapiegesprächen, von denen Video- oder Audioaufnahmen gemacht wurden. Die Aufnahmen wurden teilweise an der abteilungseigenen psychotherapeutischen Praxisstelle gemacht (*Universität Zürich, Klinische Psychologie, Psychotherapie und Psychoanalyse*), teilweise wurden sie uns freundlicherweise von der *Ulmer Textbank* (Mergenthaler & Kächele, 1994) zur Verfügung gestellt. Die verwendeten Transkripte stammen aus verschiedenen Therapien, einerseits aus der 326-stündigen psychotherapeutischen Behandlung einer Klientin mit dem Pseudonym *Wilma* (Boothe et al., 2010; Brändle, 2008; Jung, 2009; Stärk, 2008). Als Datengrundlage stehen mit wenigen Ausnahmen sowohl alle Transkripte der 326 Therapiestunden zur Verfügung als auch die Erzählungen (mehr als 800 Erzählungen), die im Rahmen von Praktika und Tutoraten an der Abteilung extrahiert wurden. Daneben werden zwei Transkriptsammlungen verwendet, die von der Ulmer Textbank zur Verfügung gestellt wurden. Dies betrifft einerseits Transkripte der Therapie des Patienten *Gustav*, der eine psychoanalytisch orientierte Therapie von 388 Stunden absolvierte; transkribiert sind davon allerdings nur 53 Stunden, die in der vorliegenden Studie verwendet werden (vgl. Haase, 2010; Rieser, 2010). Ebenfalls von der Ulmer Textbank stammen Transkripte einer psychoanalytischen Behandlung von 517 Stunden der Patientin *Amalie*, von denen 201 Stunden in transkribierter Form vorliegen. Dieses Datenmaterial ist Gegenstand von zahlreichen Psychotherapie-Studien (Albani et al., 2008). Sämtliche Patienten haben ihre Einwilligung für die wissenschaftliche Verwendung der Daten gegeben.

Die Video- bzw. Audioaufnahmen wurden zum grössten Teil nach den Regeln der Ulmer Textbank (Mergenthaler, 1992) transkribiert.¹ Eine grundsätzliche Unterscheidung für die vorliegende Arbeit betrifft die Analyse von Gesprächsausschnitten aus Transkripten (ganze Gespräche werden ausschnittsweise untersucht) gegenüber der Analyse von Erzählungen (Narrative gem. Erzählanalyse JAKOB, small stories), die eine besondere Gesprächsform darstellen und nur einen Teil des Textmaterials repräsentieren. Um einzelne Phänomene nicht nur im Kontext von Therapiegesprächen darzustellen, sondern ihr Vorkommen auch im «allgemeinen» Sprachgebrauch zu dokumentieren und mit der Verwendung im Therapiekontext vergleichen zu können, wurden ausserdem öffentlich zugängliche und durchsuchbare Textkorpora (gesprochener) deutscher Sprache verwendet, einerseits aus dem *Archiv für gesprochenes Deutsch* am IDS (Institut für Deutsche Sprache, Mannheim), Korpora aus dem *DWDS* (Digitales Wörterbuch der Deutschen Sprache) und ein deutschsprachiges Korpus, das in der *sketch engine* (siehe Kapitel 8.3.2) enthalten ist.

Die aktuelle Grösse des JAKOB-Lexikons (ca. 7000 Einträge) bringt es mit sich, dass die Beschreibung der Konstruktionen nur musterhaft oder prototypisch erfolgen kann. Die in der vorliegenden Arbeit verwendeten Beispielkonstruktionen sind online im Lexikon zugänglich.

8.1 Gesprächsdaten, Korpora

Allgemeine Überlegungen zu *Gesprächskorpora* und zu den Methoden der Korpuslinguistik wurden bereits im Kap. 5 vorgestellt. Die Verfügbarkeit von grossen, für elektronische Verwendung aufbereiteten Korpora war eine Voraussetzung für die Entwicklung von geeigneten Werkzeugen und Prozeduren, um relevante und signifikante Sprach- und Wortmuster aus den Korpora extrahieren zu können. Die Erforschung von linguistischen Fragestellungen setzt Korpora voraus, die sowohl mit einem *Tagger* mit morphosyntaktischen Informationen versehen als auch mit einem *Parser* mindestens ansatzweise mit syntaktischen Funktionen angereichert wurden (Duffner & Näf, 2006, S. 16).

¹Transkriptionsregeln im Anhang.

8.1.1 Liste der verwendeten Korpora und praktisches Vorgehen

Neben den bereits erwähnten drei Korpora mit Gesprächen aus Psychotherapien (Wilma, Gustav, Amalie) werden weitere öffentlich verfügbare Korpora für Vergleichszwecke herangezogen. In erster Linie sind dies Gesprächsdaten aus dem *Archiv für Gesprochenes Deutsch* (AGD), das transkribierte Gesprächsdaten archiviert hat und sie für Forschungszwecke und für die weitere wissenschaftliche Auswertung zur Verfügung stellt. Ein grosser Teil der Psychotherapie-Korpora und der JAKOB Erzählenden stammt aus dem Deutschschweizer Sprachraum, weshalb zusätzlich das für wissenschaftliche Zwecke öffentliche *Schweizer-Textkorpus* (CHTK) in die Recherche mit einbezogen wird. In der *sketch engine*² integriert und öffentlich zugänglich ist das *Textkorpus deWAC* (Baroni & Kilgariff, 2006), das über das Internet zugängliche deutsche Textdaten enthält. Schliesslich werden teilweise auch die in *COSMAS* (Corpus Search, Management and Analysis System) verfügbaren Korpora im Archiv der geschriebenen Sprache zu Vergleichszwecken herangezogen. Damit wird das Problem der *Gesprächssortenbeschränkung* auf die Gespräche aus Psychotherapien etwas entschärft. Die Gespräche aus dem Schweizer Sprachraum wurden ins Standarddeutsche transkribiert, allerdings unter möglichst weitgehender Beibehaltung des besonderen Wortschatzes. Einschränkungen sind bei den entsprechenden Korpora aufgeführt. Es folgt eine kurze Vorstellung der verwendeten Korpora.

8.1.2 WIL: Korpus Wilma

Das Korpus WIL besteht aus den Transkripten einer psychoanalytischen Therapie, die in den 1990er Jahren an der Praxisstelle der Abteilung Klinische Psychologie, Psychotherapie und Psychoanalyse durchgeführt wurde und sich über sechseinhalb Jahre und 326 Stunden erstreckte. Die Sitzungen wurden grösstenteils aufgezeichnet und nach den Transkriptionsregeln der Ulmer Textbank (Mergenthaler, 1992) anonymisiert und transkribiert. Bei der Klientin *Wilma* handelt es sich um das Pseudonym für eine etwa 40-jährige Frau, welche die Praxisstelle aufsuchte, um nach eigenen Angaben ihre Probleme im Umgang mit Männern zu bewältigen.

Die Transkription ist teilweise mangelhaft; diverse Probleme ergeben sich aus versuchten Übersetzungen und Nichtverstehen des Dialekts. Neben den

²siehe Kapitel (8.3.2)

Transkripten, die für die 326 Stunden bis auf kleine Ausnahmen zur Verfügung stehen, sind auch alle Erzählungen von Wilma erfasst (über 800 Erzählungen gemäss Kriterien der Erzählanalyse JAKOB). Aus der erzähl-analytischen Forschung an der Abteilung liegen mehrere Studien und Publikationen zur Patientin Wilma vor (Boothe et al., 2010; Brändle, 2009; Grimmer, 2000; Jung, 2009; Stärk, 2009).

Die Tabelle (8.1) fasst die statistischen Daten zum Korpus WIL zusammen:³

Tabelle 8.1: Transkripte Wilma (WIL)

Spaltentitel	Anzahl	Bem
Transkripte	326 Sitzungen	Therapiedauer: 326 Sitzungen
Erzählungen	886	(in JAKOB-Projekten erfasst)
Wortfrequenz	1'925'586	(Angabe aus MaxQDA)
Tokens	2'424'817	(Angabe aus Sketch Engine)
Verwendung <ich>	72435	(sketch engine)
Verben	277'248	(sketch engine)

8.1.3 AMA: Korpus Amalie

Amalie ist das Pseudonym für eine 35-jährige Frau, die ca. 1940 geboren wurde. Sie lebte zum Zeitpunkt der Therapie in Süddeutschland und ging ab ca. 1978 drei Mal wöchentlich zu psychoanalytischen Sitzungen. Die Therapie umfasste 517 Stunden und dauerte fünf Jahre (Albani et al., 2008). Von den 517 Sitzungen stehen 208 Sitzungen transkribiert für die Korpusanalyse zur Verfügung (siehe Tab. 8.2).

Tabelle 8.2: Transkripte Amalie (AMA)

Spaltentitel	Anzahl	Bem
Transkripte	208 Sitzungen	Therapiedauer: 517 Sitzungen
Erzählungen	284	(in JAKOB-Projekten)
Tokens	1'128'736	(Sketch Engine)
Verwendung <ich>	36683	(sketch engine)
Verben	123'766	(sketch engine)
Datenherkunft		Ulmer Textbank

³Tokens beinhalten sämtliche wortähnlichen Gebilde, also auch die in den Transkripten verwendeten Kürzel für Patient und Therapeut, *P* und *T*, weshalb die Tokenzahl um einiges grösser ist als die Wortzahl.

8.1.4 GUS: Korpus Gustav

Das Korpus GUS besteht aus Transkripten des Patienten *Gustav*, zur Verfügung gestellt von der Ulmer Textbank (siehe Tab. 8.3). Zu 54 seiner 390 Stunden umfassenden psychoanalytischen Behandlung liegen Verbatimprotokolle vor, die sich über den ganzen Behandlungszeitraum verteilen (Haase, 2010).

Tabelle 8.3: Transkripte Gustav (GUS)

Spaltentitel	Anzahl	Bem
Transkripte	54	Gesamtstunden: 390
Erzählungen	165	(in JAKOB-Projekten)
Tokens	297'515	(Sketch Engine)
Verwendung <ich>	10803	(sketch engine)
Verben	28097	(sketch engine)
Datenherkunft		Ulmer Textbank

8.1.5 AGD: Archiv für Gesprochenes Deutsch

Das *Archiv für Gesprochenes Deutsch* (AGD) beinhaltet in diversen Forschungsprojekten erstellte Korpora von gesprochener deutscher Sprache und stellt sie für Forschungszwecke zur Verfügung. Das Archiv und die damit verbundene *Datenbank Gesprochenes Deutsch* (DGD) sind am Institut für deutsche Sprache (IDS) in Mannheim beheimatet. Es sind mehrere Korpora verfügbar, unter anderem zu Beratungsgesprächen, zu biographischen und Reiseerzählungen, zu Talkshows, Diskussionen und Interviews im Fernsehen, aber auch zu lokalen Sprachausprägungen (z. B. Stadtsprache Mannheim) und deutschen Mundarten. Ausserdem ermöglicht es das Datenbanksystem, die Tonaufnahmen der Belegstellen sofort anzuhören.⁴

Da die Korpora des AGD teilweise klein sind, und da im vorliegenden Projekt eher seltene Konstruktionen untersucht werden, wurden immer *alle verfügbaren Korpora* in die Suchabfragen einbezogen, um genügend Treffer zu erhalten. Ein Beispiel: Die Suche nach «Karren» (siehe Kapitel 9.2.2) ergibt unter Einbezug aller Korpora 52 Treffer, bei einer Beschränkung auf Korpora zur verbalen Interaktion im sozialen Kontext ergeben sich hingegen keine Treffer.

⁴Eine Liste der Korpora mit genauen Angaben zu Charakteristika und Grösse ist hier zu finden:

<http://dsav-wiss.ids-mannheim.de/DSAv/KORPORAL.HTM> (3.11.2010)

Tabelle 8.4: Korpora aus dem AGD

Spaltentitel	Anzahl	Bem
Korpora	27	(Gesprächskorpora)
Aufnahmedauer	über 2100 Stunden	
Wörter/Tokens	über 5.7 Millionen	
öffentlich	ja	für wissenschaftliche Zwecke

8.1.6 CHTK: Schweizer Textkorporus

Das *Schweizer Textkorporus* wird am Deutschen Seminar der Universität Basel aufgebaut und gehört zu einem internationalen Forschungsprojekt (Digitales Wörterbuch der Deutschen Sprache, DWDS) zur deutschen Standardsprache des 20. Jahrhunderts. Das Korpus besteht aus verschiedenen Textsorten (Zeitungsartikel, Werbung, Formulare, Anleitungen, Ratgeber, populäre Fachliteratur, Jugend- und Trivalliteratur, Belletristik etc.). Das Gesamtkorpus trägt den Namen *C4* und umfasst rund 80 Millionen Textwörter.⁵ Das Schweizer Textkorporus ist ein Teilprojekt von *C4*, vereint deutschsprachige Texte von Schweizer Autoren aus dem 20. Jahrhundert, «stellt eine ausgewogene Repräsentation des Deutschschweizer Wortschatzes dar und kann als Grundlage für spezifisch schweizerische lexikographische Bedürfnisse dienen» (Projektbeschreibung CHTK).⁶ Dieses Korpus beinhaltet nur geschriebene Sprache.

Zum Vergleich: Die oben beim AGD erwähnte Recherche nach dem Lemma «Karren» ergibt im CHTK 152 Treffer.

Tabelle 8.5: Schweizer Text Korpus (CHTK)

Spaltentitel	Anzahl	Bem
Texte	ca. 12'000	(ausgewogen nach div. Kriterien)
Textwörter	ca. 20 Millionen	(Angabe CHTK)
öffentlich	ja	(mit Registrierung)

⁵siehe URL: <http://beta.dwds.de/resource/c4/> (3.11.2010)

⁶siehe URL: <http://www.dwds.ch/> und <http://chtk.unibas.ch/search> (4.11.2010)

8.1.7 deWAC: Deutschsprachiges Korpus der Sketch Engine

Die im Kapitel (8.3.2) näher beschriebene *sketch engine* ist eine Internetapplikation, die geeignete Prozeduren für die Analyse von grossen Korpora bereitstellt, aber auch selber Korpora für verschiedene Sprachen online zur Verfügung stellt. Für die deutsche Sprache wurde das Korpus *deWAC*⁷, eine Sammlung von deutschen Texten aus dem Internet mit über 1.5 Billionen Tokens aufbereitet. Auch hier handelt es sich also um ein Korpus geschriebener Sprache. Die Texte wurden durch ein spezielles, *web crawl* benanntes Verfahren (Baroni & Kilgariff, 2006) direkt aus dem Internet kompiliert, mit part-of-speech tags versehen und lemmatisiert. Word sketches stehen für dieses Korpus aktuell noch nicht zur Verfügung, weil die Texte nicht hinreichend getaggt sind (gemäss Angaben auf dem Internet: «word sketches currently in preparation».⁸

Das deWAC bietet dank Lemmatisierung und Wortarten-Tags sehr gute und komfortable Suchmöglichkeiten. Ausserdem steht die Funktion *collocations* zur Verfügung. Damit werden zu den gefundenen Ausdrücken die besten Kollokationskandidaten angezeigt, basierend auf verschiedenen statistischen Assoziationsmassen.

Die Suche nach dem Wort «Karren» liefert für dieses grosse Korpus 3598 Treffer.

Tabelle 8.6: Korpus deWAC (deWAC)

Spaltentitel	Anzahl	Bem
Korpus	Texte aus dem Internet	(Div. Textsorten)
Tokens	1,627,169,557	(gemäss sketch engine)
öffentlich	ja	

8.1.8 Weitere Korpora

Es gibt deutsche *Korpusdatenbanken*, die eine Mittelstellung zwischen Korpus und elektronischem Wörterbuch einnehmen, so zum Beispiel die digitalen Wörterbuchsysteme *Digitales Wörterbuch der Deutschen Sprache* (DWDS) und *Wortschatz Leipzig*. Diese sind wie Wörterbücher aufgebaut, bieten aber zu jedem Eintrag Belegstellen an, die online eingesehen werden

⁷URL: <http://the.sketchengine.co.uk/auth/corpora/> (4.11.2010)

⁸siehe URL: <http://trac.sketchengine.co.uk/wiki/Corpora/DeWaC> (4.11.2010)

können. Das *TIGER-Korpus* ist eine Baumbank für das Deutsche und besteht aus Zeitungstexten der Frankfurter Rundschau. Eine Baumbank ist eine Sammlung von Sätzen, die mit syntaktischen Informationen (Phrasenstruktur und grammatische Funktionen) sowie Wortklassen, Morphologie und Lemma annotiert sind.

Allgemein ist zu beklagen, dass für *gesprochene Sprache* nur sehr wenig Korpusdaten öffentlich zugänglich sind. Einerseits ist die öffentliche Zugänglichkeit bei Korpora gesprochener Sprache ein grösseres Problem als bei Korpora geschriebener Sprache, weil Gespräche naturgemäss im kleineren, in der Regel nicht öffentlichen Rahmen ablaufen. Auf der andern Seite ist auch das Problem der Aufbereitung nicht zu unterschätzen. Gespräche müssen vor der Erfassung in einem elektronisch aufbereiteten und annotierten Korpus in zeitaufwendiger Kleinarbeit transkribiert werden.

8.2 Lexika und MWE-Sammlungen

Für die Beschreibung und Annotierung der prototypischen Lexikon-Beispiele sowohl im Durchführungsteil als auch im überarbeiteten Lexikon wurden Beschreibungen aus verschiedenen lexikalischen Ressourcen miteinander und mit den Korpusdaten verglichen, um möglichst vielfältige und repräsentative Daten als Grundlage für die Lexikoneinträge zu haben. Folgende Online- und Print-Wörterbücher wurden und werden dafür verwendet:⁹

- *Der Kleine Wahrig* (Wahrig, 2007). Dieses Wörterbuch hat für die Erzählanalyse JAKOB eine längere Tradition, weil die Verben im JAKOB-Lexikon seit dem Jahre 1999 mit dem *Satzmuster nach Wahrig* annotiert werden.
- *Der Duden 11* (Scholze-Stubenrecht & Wermke, 2008) ist ein bewährtes Nachschlagewerk für deutsche Redewendungen.
- *Das Valenzwörterbuch deutscher Verben* (VALBU) (Schumacher et al., 2004) wurde am IDS Mannheim entwickelt und bietet eine reichhaltige syntaktische und semantische Beschreibung von 638 deutschen Verben mit ihrer Valenzumgebung. Das Wörterbuch ist als E-VALBU auch online zugänglich.¹⁰

⁹Diese Aufzählung ist sehr heterogen und umfasst neben ausgearbeiteten und auf Vollständigkeit ausgerichteten Nachschlagewerken auch kleinere Projekte, die meiner Ansicht nach innovative Wege bei der Beschreibung von Lexikoneinträgen beschreiten.

¹⁰siehe URL: <http://hypermedia2.ids-mannheim.de/evalbu/> (4.11.2010)

- *Ideomatisierte verbale Phraseologismen* (IVPs) im Deutschen ohne freie Lesart: Söhn (2006) präsentiert eine erklärende Beschreibung von 190 ideomatisierten verbalen Phraseologismen.
- *Phraseo-Lex*: Auch Keil (1997) stellt eine Beschreibung und Bewertung *verbaler Phraseologismen* und ihrer spezifischen Eigenschaften auf linguistischem und computerlinguistischem Hintergrund vor. Die Semantik verbaler Phraseologismen steht im Vordergrund. Die lexikalische Datenbank Phraseo-Lex dient als prototypisches Werkzeug zur systematischen Erfassung und Klassifizierung verbaler Phraseologismen.
- *Dornseiff: Deutscher Wortschatz nach Sachgruppen* (Dornseiff, 2004). Der Dornseiff ist ein onomasiologisches Wörterbuch und bildet den deutschen Wortschatz nach Sachgruppen geordnet ab. Jeder Eintrag präsentiert Wörter (und Ausdrücke), die zu einer Sache oder einem Begriff gehören. Seit der 8. Auflage (2004) erfolgt die Aufnahme der Wörter in bestimmte Gruppen korpusbasiert.
- *FrameNet* (Ruppenhofer et al., 2010). Das *Berkeley FrameNet Projekt* ist eine englischsprachige lexikalische Ressource, die auf der *Frame-semantik* basiert und deren Struktur aus Korpusevidenzen aufgebaut ist.¹¹

Ausserdem gibt es weitere *Nachschlagewerke*, die sich auf Verben beziehen und Top-Down Typologien für Verben entwerfen, die zwar interessant, aber für eine praktisch nützliche Kategorisierung im Lexikon oft wenig hilfreich sind (z. B. Ballmer & Brennenstuhl, 1986; Harras et al., 2004, 2007; Levin, 1993; Pottelberge, 2001).

Schliesslich kann auch eine Google-Recherche im Internet über bestimmte, spezielle oder seltene Ausdrücke sehr brauchbare Hinweise zu Bedeutung und Sprachgebrauch geben.

8.3 Methoden

Ziel der vorliegenden Studie ist es, in den vorgestellten Korpora aus Therapiegesprächen prototypische Beispiele von Ausdrücken zu finden, die gesprächs- und kontextspezifische Bedeutung transportieren, diese als Lexikoneinträge mit einem reichhaltigen Set von lexikalisch nützlichen Eigen-

¹¹siehe URL: <http://framenet.icsi.berkeley.edu/> (4.11.2010)

schaften zu versehen mit dem Ziel, die praktische Verwendbarkeit eines solchen Lexikons für das Verstehen (die Bedeutungsrezeption) von Gesprächsausschnitten zu prüfen. Dabei ist einerseits ein gewisser Kontext vorgegeben (Genre Psychotherapiegespräch), und andererseits werden ausgehend von einzelnen Konstruktionen lokale Kontexte und Funktionszusammenhänge gesucht, in denen diese Konstruktionen Verwendung finden (Günthner & Bucker, 2009, S. 4).

Am Anfang steht eine gesprächsanalytische Untersuchung von einzelnen Gesprächsstellen und auffallenden sprachlichen Phänomenen (Deppermann, 2006c, S. 52). Wie wurden diese gefunden? Ein systematisches Vorgehen ist für die grossen Textmengen und im Rahmen dieser Arbeit nicht möglich. Es wurden deshalb einzelne Gesprächsstellen gesucht, die auf Grund der bestehenden automatisierten und manuellen Kodierprozeduren erfahrungsgemäss nicht adäquat kodiert werden können, d. h. wo die Bedeutung, die aus dem Kontext und der Situation konstituiert wird, nicht aus den lexikalischen Grundlagen eruiert werden kann. Diese Ausdrücke sollen mit einem neuen, auf Konstruktionen beruhenden Lexikon besser kodiert werden können. Anschliessend an die Analyse einer einzelnen Textpassage erfolgt der Vergleich mit weiteren Textstellen mit gleichen oder ähnlichen Konstruktionen und die Erweiterung der gesprächsanalytischen Befunde mit Korpusmethoden. Schliesslich sollen neue Lexikoneinträge mit den passenden Merkmalen versehen werden. Folgende Fragen sollen dabei beantwortet werden (siehe auch Einleitung):

1. Können *Konstruktionen* (Lexikoneinträge) mit den vorgeschlagenen Methoden aus den Korpusbeispielen erschlossen und adäquat dargestellt werden, und welches sind die Anforderungen dafür?
2. Können *gleich lautende oder ähnliche Konstruktionen* mit unterschiedlicher Bedeutung (Homonymie und Polysemie) im Lexikon so differenziert dargestellt werden, dass eine Unterscheidung auf Grund der Lexikonmerkmale möglich ist?
3. Repräsentieren Lexikoneinträge eine *Grundbedeutung* (Kernsemantik), die als Grundlage für die *Bedeutungskonstitution* in der Gesprächssituation verwendet werden kann?
4. Wieviel *Kontext* müssen Lexikoneinträge enthalten? Wie dekontextualisiert dürfen sie sein?
5. Ist der Begriff der *Konstruktion* ein passendes Konzept für die Abbildung lexikalischer Phänomene?

6. Wie funktioniert die *Bedeutungskonstitution dynamisch* in der Interaktion? Können über diesen Prozess Erkenntnisse gefunden und als Einsichten oder Hypothesen für die Gesprächs- und Erzählanalyse gewonnen werden?
7. Kann die *automatische Kodierung* in der Erzählanalyse JAKOB mit den geplanten Lexikonergänzungen verbessert werden?

8.3.1 Gesprächsanalyse

Die Auswahl der zu untersuchenden Konstruktionen in den Korpora erfolgt mit Analysen von kleinen Gesprächsausschnitten, bei Deppermann (2006c, S. 52) als mikroskopischer Einstieg bezeichnet (gegenüber einem Zugang zu den Daten, der zuerst die grossen Strukturen des Gesprächs entwickelt). Auffällige Phänomene in der Sprache der Klienten in Hinsicht auf die Verwendung von Konstruktionen und deren Erfassung und Disambiguierung im Lexikon werden als Beispiele und typische Fälle aufgenommen, weil sie als Ausgangspunkt für die erwähnten Forschungsfragen interessant scheinen (Deppermann, 2006c, S. 52). Für die ausgewählten Konstruktionen wird eine *Kollektion* aller passenden Stellen in den Transkripten zusammengestellt (Deppermann, 2006c, S. 37); diese Kollektion fällt für die vorliegende Untersuchung weitgehend mit einer bereinigten Konkordanzliste (KWIC) aus den verwendeten Transkripten zusammen.¹² Dabei sollen folgende Fragen beantwortet werden (vgl. Deppermann, 2006c, S. 13–15):

- Wieso wird dieser *Gesprächsausschnitt* zitiert? Was an der Konstruktion wird untersucht?
- Welche *Kontexte* und *Inhalte* sind typisch für das Beispiel? Allenfalls ist eine Untersuchung von weiteren transkribierten Gesprächsstellen nötig, oder die weiteren Stellen werden pauschal erwähnt und kategorisiert.
- *Form-* und *funktionsbestimmte* Analyse: mit welchen Mitteln und warum werden Sinn und Ordnung im Gespräch hergestellt?
- *Generalisierungen*: Wie ist der Sprachgebrauch in den Therapiegesprächen im Vergleich zu den verfügbaren Referenzkorpora? Kann man daraus Regeln formulieren?

¹²Die *sketch engine* erlaubt jederzeit von der zeilenweise knappen Darstellung der Kookkurrenzen in der KWIC-Liste in eine Ansicht zu wechseln, in welcher die Konstruktionen innerhalb eines grösseren Textfensters angezeigt werden.

- Ziel ist eine *Abbildung* der Konstruktion und des Konstruktionsverhaltens mit Hilfe der vorhandenen und neuer Lexikonmerkmale. Welche Erkenntnisse können für das Lexikon übernommen werden?
- Kann eine (*Kern-*)*Bedeutung* im Lexikon festgeschrieben werden? Wie können Emergenzprozesse berücksichtigt und beschrieben werden? Wie soll mit der dynamischen Bedeutungskonstitution umgegangen werden?

Die Analyse der Texte wird angelehnt an das bei Deppermann (2006c, S. 55) beschriebene gesprächsanalytische Verfahren durchgeführt. Es folgt eine kurze Darstellung des idealtypischen Vorgehens einer *Detaillierten Sequenzanalyse am Einzelfall* (Deppermann, 2006c, S. 55 ff.). Diese Vorgaben einer umfassenden Gesprächsanalyse sind nicht sakrosankt, sondern müssen den Untersuchungsinteressen und oft auch den Gegebenheiten des Datenmaterials angepasst werden. Für das vorliegende Projekt sind nicht alle Schritte gleich relevant und werden deshalb teilweise modifiziert und angepasst.

- I *Paraphrase und Handlungsbeschreibung*: Die Gesprächsstelle wird inhaltlich paraphrasiert und die vorliegenden sprachlichen Handlungen werden beschrieben. Die weitere Arbeit sollte sich auf das Originalmaterial stützen, nicht auf die Paraphrase.¹³
- II *Äußerungsgestaltung und Formulierungsdynamik*: Die Art und Weise, wie gesprochen wird mit den Parametern Phonetik, Prosodie, Grammatik (Wortfolge, syntaktische Konstruktionen), lexikalische Wahlen und Stilistik (Routineformeln) werden beschrieben. Gibt es Kontextualisierungshinweise (Elemente, die implizit einen Kontext einführen)?
- III *Timing*: Dazu gehören der zeitliche Ablauf der Formulierungen verschiedener Sprecher und weitere Merkmale wie Sprecherwechsel, Rede-recht, Koordination des Gesprächs.
- IV *Kontextanalyse*: In welchem Kontext läuft der untersuchte Gesprächsausschnitt (die fokale Äußerung) ab? Wie hat sich die Gesprächssituation zum aktuellen Moment hin entwickelt? Kontexte umfassen Vorwissen, Fähigkeiten der Sprecher, Rechte, Pflichten, Ereignisse etc. Zum Verständnis genügt nicht grammatisch-semantisches Wissen, sondern es ist dafür Handlungs- und Weltwissen erforderlich.

¹³Deppermann (2006c) gibt zu jedem Schritt der Sequenzanalyse einen detaillierten Fragenkatalog vor.

- V *Folgerwartungen*: Jede Äusserung orientiert sich am vorausgehenden Kontext und bildet selber den Kontext für die folgenden Äusserungen. Kontext, Redebeiträge und soziale Erwartungen eröffnen Leerstellen, die durch weitere Redebeiträge gefüllt werden können (responsive vs. projektive Eigenschaften, vgl. Linell, 2009, S. 302).
- VI *Interaktive Konsequenzen*: Durch die Interpretation der vorausgegangenen Gesprächszüge werden mögliche *Bedeutungen* von Gesprächsbeiträgen ersichtlich, die sich durch Vagheit, Flexibilität, Revidierbarkeit und Aushandelbarkeit auszeichnen. Diese Phänomene zeigen sich in Reformulierungen (formulations, vgl. Linell, 2009, S. 303), Reframings (vorangegangene Äusserungen in einen neuen Rahmen stellen), Rechtfertigungen (accounts) und Korrekturen (repair). Interaktive Konsequenzen zeigen sich in den so genannten *3 Positionen* (First position: fokaler Beitrag des ersten Sprechers; second position: unmittelbar folgende Reaktion des Gesprächspartners; third position: Reaktion des ersten Sprechers auf die Reaktion des Gesprächspartners).
- VII *Sequenzmuster und Makroprozesse*: Verschieden aufgebaute Sequenzen (z. B. Vorbemerkungen oder Einschübe) und verschiedene Abläufe strukturieren grössere Gesprächseinheiten. Die Abfolge und die Zusammenhänge von Beitragspositionen werden als Makroprozesse bezeichnet. Damit wird der Fokus der Gesprächsanalyse von einzelnen Redebeiträgen auf die Diskursebene erweitert.

Die in den Analyseschritten gefundenen *Gesprächspraktiken* müssen sowohl in Bezug auf ihre prozeduralen Charakteristika (Wie-Seite) als auch in Bezug auf ihre funktionalen Eigenschaften (Wozu-Seite) untersucht werden; das Wie und das Wozu muss in einen systematischen Zusammenhang gebracht werden (Deppermann, 2006c, S. 79). Schliesslich müssen auch Wissensressourcen von ausserhalb des Gesprächs hinzugezogen und ihre Auswirkungen auf den Gesprächsprozess geprüft werden.

Als methodologische Leitlinien für den Gesprächsanalytiker gelten gemäss Deppermann (2006c): «seine Aussagen so weit wie möglich auf die Aktivitäten der Gesprächsteilnehmer zu stützen», «zu zeigen, dass seine Interpretationen mit den Details des interaktiven Geschehens lückenlos vereinbar sind» und «zu explizieren, dass und wie die Äusserungen der Gesprächsteilnehmer im Sinne seiner Aussagen interpretiert werden können» (S. 51). Damit wird wieder das schon mehrmals angesprochene inhärente Problem der Dekontextualisierung und Abstrahierung von sprachlichen Phänomenen

in einem Lexikon angesprochen, wie dies auch Günthner und Bückler (2009) formulieren:

Methodologisch hat dies die Konsequenz, dass sprachliche Phänomene nicht beliebig dekontextualisiert werden dürfen, sondern immer im jeweiligen konkreten Interaktionszusammenhang zu beschreiben sind. Darüber hinaus dürfen weder die Medialität und Handlungsbezogenheit sprachlicher Phänomene noch ihre inhärente Dialogizität, ihre Sequenzialität und die Zeitlichkeit ihrer Entfaltung ausgeblendet werden. (S. 4)

*MaxQDA*¹⁴, ein computerunterstütztes Inhaltsanalyseprogramm, wird im vorliegenden Projekt als Arbeitshilfe für die Gesprächsanalyse verwendet. Das Programm erlaubt die Kommentierung und Kodierung von Textstellen und die effiziente Verwaltung dieser Textstellen auf übersichtliche Art. Für weitergehende korpuslinguistische Fragestellungen (Häufigkeiten, Kollokations- und Kookkurrenzanalysen) ist diese Applikation weniger geeignet; dafür wird die im folgenden Kapitel beschriebene *sketch engine* verwendet.

8.3.2 Korpusanalyse und Sketch Engine

Nachdem die gewählten Gesprächsausschnitte mit gesprächs- und erzähl-analytischen Methoden untersucht wurden, werden die aufgestellten Vermutungen und Hypothesen an den Korpusdaten überprüft, was möglicherweise eine Generalisierung erlaubt, die allein mit dem selektiven gesprächs-analytischen Zugang nicht möglich ist. Folgende Fragen sollen damit beantwortet werden: Ist die Konstruktion X für die Einzelperson, das Gespräch, die ganze Therapie typisch, wird damit immer wieder das Gleiche ausgedrückt? Wie lautet das Konzept, das damit ausgedrückt wird? Kommt die Konstruktion überhaupt so häufig vor, dass ihr eine spezifische Bedeutung zugeordnet werden kann (diese Frage kann auf Grund der kleinen Korpora nur bedingt beantwortet werden)? Ist die Konstruktion lexikalisiert, teillexikalisiert oder eher schematisch? Dabei ist die konversationsanalytische Bedingung zu beachten, dass die Generalisierung auf den gefundenen Teilnehmer- und Sprecherkategorien beruht und nicht von aussen herangebracht wird (Fischer, 2007).

Für die korpuslinguistische Analyse wird die *sketch engine* verwendet (Kilgariff, Rychlý, Smrz & Tugwell, 2004),¹⁵ ein Korpusanalysewerkzeug,

¹⁴URL: <http://www.maxqda.de/> (15.11.2010)

¹⁵URL: <http://the.sketchengine.co.uk/> (1.11.2010)

mit dem *Konkordanzen* erstellt werden können, die das Auftreten eines gesuchten Wortes/Ausdrucks im Korpus abbilden, in der Regel zeilenweise mit dem Suchwort im Zentrum und einem kurzen Textausschnitt vor und nach dem Suchwort (*KWIC* - keyword in context). Dieses Textfenster kann nach Belieben vorwärts und rückwärts erweitert werden, um den Kotext und Kontext einer Äußerung darzustellen. Die *sketch engine* ermöglicht neben der Analyse des eingebauten deutschen Korpus *deWaC* (siehe Kapitel 8.1.7) den Import und die Bearbeitung eigener Korpora. Für die korpuslinguistischen Arbeitsschritte werden die Transkripte in die *sketch engine* geladen, und die Texte werden lemmatisiert und mit Wortarten-Tags versehen. Aus dem ursprünglichen Text wird ein so genanntes *vertikales File* erstellt: Jedes Token wird einzeln auf eine neue Zeile geschrieben, die Zeile beinhaltet die Angaben zu Wortform, Lemma und Wortart. Dies passiert mit annehmbarer Geschwindigkeit und vor allem mit einer ebenfalls annehmbaren Präzision. Obwohl für englische Sprache entwickelt, ist die *sketch engine* mit Hilfe geeigneter Regeln für das Tagging und Parsing auch für deutschsprachige Texte verwendbar.

Mit dieser Aufbereitung können *Konkordanzen* und *Kookurrenzanalysen* erstellt werden. SKE generiert auch einen Thesaurus und ermöglicht es, Kollokationsdifferenzen aufzuspüren, die Ähnlichkeiten und Differenzen im Kollokationsverhalten verschiedener Ausdrücke aufzeigen können (Kilgariff et al., 2004). So können z. B. zu einem Verb alle Kollokationskandidaten lemmatisiert und mit der Häufigkeit ihres Auftretens ausgegeben werden, zusätzlich mit einem statistischen Assoziationsmass für die Signifikanz des gemeinsamen Auftretens versehen (unter anderen *mutual information* und *logDice-Score*, siehe Kapitel 5.1).

Die eigentliche Spezialität der *sketch engine* sind *word sketches*, die detailliert Aufschluss über das grammatikalische und kollokationale Verhalten eines Wortes/Ausdrucks in einem bestimmten Korpus geben. Um das Verhalten der Einzelwörter genauer zu bestimmen, genügt es nicht, dass die einzelnen Lexeme mit Wortarten-Tags versehen sind, sondern es braucht zusätzlich eine einfache Grammatik, um das Verhalten der Lexeme im Kontext zu bestimmen (Ivanova, Heid, Schulte Im Walde, Kilgariff & Pomikalek, 2008). Eine *sketch grammar* ist ein auf regulären Ausdrücken basierendes Regelwerk, das die syntaktischen Beziehungen in den Korpussegmenten über den part-of-speech tags regelt (shallow parsing).

Falls zu den einzelnen Wortformen genügend linguistische Informationen vorhanden sind, können mit *word sketches* auch konstruktionsähnliche Gebilde gefunden werden, z. B. alle Vorkommen eines bestimmten Verbs gemeinsam mit einem bestimmten Nomen innerhalb eines kleinen Textfens-

ters. Dies erfordert allerdings eine ausgefeilte *sketch grammar* und setzt voraus, dass die Tokens auch morphologische Informationen (zusätzlich zur Wortart) beinhalten. Da sich verschiedene Satzkonstruktionen und Verbpat-terns im Deutschen hauptsächlich durch die Kasusmarkierung unterscheiden (Wortformen der Nominalphrasen bezeichnen den Kasus) und nicht durch ihre Reihenfolge, ist eine *sketch grammar* für das Deutsche wesentlich komplizierter zu erstellen als eine solche für Englisch, bei der stark auf die *Reihenfolge* der Konstituenten gebaut werden kann. Ivanova et al. (2008) präsentieren eine sehr weit entwickelte und komplizierte *sketch grammar* für das Deutsche, die allerdings schwierig zu handhaben ist und grosse Performanceprobleme verursacht.¹⁶ Da zur Zeit keine einsatzfähige schlanke *sketch grammar* für das Deutsche vorhanden ist, wurde für das vorliegende Projekt eine rudimentäre Grammatik zusammengestellt, die einfache Zusammenhänge von Verben, Nomen und Adjektiven aufzeigen kann.

8.3.3 Corpus Pattern Analysis

Mit Hilfe der Korpusanalyse werden die Verbpatterns im Sinne der *Corpus Pattern Analysis* (CPA) erschlossen. In diesem Sinne kann auch die im Kapitel (6.11) beschriebene CPA als Methode betrachtet werden. Im Unterschied zur Konstruktionsgrammatik ist die Methode im Falle der CPA genauer festgelegt. Im vorliegenden Lexikonprojekt wird die CPA so angewendet, wie sie von Hanks und anderen für die englische Sprache beschrieben wird (Hanks, 2004; Hanks & Pustejovsky, 2005; Hanks, 2008a; Pustejovsky et al., 2004).

8.4 Fazit zum Kapitel Daten und Methoden

Die Datengrundlage für das vorliegende Projekt ist nicht als optimal zu bezeichnen. Besonders für die Durchführung der Korpusanalysen über die Gesprächskorpora wäre es vorteilhaft, mehr Daten zur Verfügung zu haben. Dass es so wenig elektronisch verfügbare und transkribierte Gesprächsdaten gibt, ist auf verschiedene Faktoren zurückzuführen. Erstens finden Gespräche meistens in einem kleineren oder in einem privaten Rahmen statt und sind entsprechend vergänglich; eine Aufzeichnung der Gesprächsdaten muss explizit geplant und meistens in einem institutionellen Rahmen durchgeführt werden. Bestimmte Gesprächstypen, wie z. B. Therapiegespräche können nicht aufgezeichnet und ausgewertet werden, wenn die Betei-

¹⁶(Persönliche Mitteilung Jan Pomikalek)

ligten nicht ausdrücklich ihre Einwilligung dazu geben. Und schliesslich ist ein wesentlicher Grund für die spärliche Datenlage darin zu sehen, dass die Transkription von aufgezeichneten Gesprächsdaten eine ausserordentlich zeitaufwendige und damit auch kostspielige Angelegenheit ist. Diese Gründe führen dazu, dass transkribierte Gesprächskorpora im Vergleich zu elektronisch verfügbaren schriftlichen Texten stark untervertreten sind. Ob schon in der vorliegenden Arbeit Gesprächsphänomene untersucht werden, ist der Anteil der Gesprächsdaten in den verwendeten Korpora klein. Ausserdem ist die Zusammensetzung des Datensamples nicht ausgewogen; die Hauptkategorien Therapiegespräch, Internettexte und Belletristik stellen eine eher willkürliche Auswahl dar.

In der praktischen Arbeit hat sich gezeigt, dass das Korpus WIL eine annehmbare Grösse hat, die den vorliegenden Aufgaben gerecht wird. Das Korpus WIL ist nach meiner Ansicht über einer kritischen Grösse, die notwendig ist, um korpusanalytische Erkenntnisse zu gewinnen, während die Korpora AMA und vor allem GUS unter dieser kritischen Grösse liegen. Auf Grund der vorliegenden Daten könnte diese kritische Grösse auf ca. 2 Mio Tokens veranschlagt werden.¹⁷ Dass einzelne sprachliche Phänomene (Phraseologismen und Idiome) trotzdem nur ein- bis zweimal vorkommen, entspricht der Empirie, widerspricht aber einer intuitiven Überbewertung salienter Phänomene.

Die Kombination von *gesprächsanalytischen* und *korpuslinguistischen* Methoden zur Erkenntnisgewinnung scheint mir ideal zu sein. Die beiden Verfahren ergänzen sich hervorragend, indem die Gesprächsanalyse mit Aspekten der Generalisierung von lokalen Eigenarten ergänzt wird, und weil die Methoden der Korpuslinguistik von den Nachteilen befreit werden, die eine rein formalistische und mechanische Anwendung von Distributionsanalysen mit sich bringen. Für die weitere Entwicklung des JAKOB-Lexikons ist die beschriebene Methodenkombination sinnvoll und sollte weiter geführt werden.

Die *sketch engine* erweist sich als ein unverzichtbares Werkzeug für die effiziente und einfach durchzuführende linguistische Analyse eigener Korpora. Öffentlich zur Verfügung stehende Korpusanalysewerkzeuge können meistens nur mit den eingebauten Korpora verwendet werden.

¹⁷Voraussetzung für diese Überlegung sind möglichst homogene Gesprächsdaten, z. B. Daten einer bestimmten Person, eines bestimmten Gesprächstyps oder eines bestimmten Dialekts.

9 Durchführung - Lexikonerstellung

Das folgende Kapitel beschreibt die aus den bisher dargestellten Konzepten und theoretischen Überlegungen folgenden neuen Anforderungen an das JAKOB-Lexikon und die Möglichkeiten der praktischen Umsetzung und Implementierung. Ausgehend von der bisherigen Form des JAKOB-Lexikons werden zuerst die bestehenden Strukturen dargestellt, dann die Neuerungen, die durch die Umstellung auf das OLIF-Format eingeführt wurden, und schliesslich die aktuellen Ergänzungen in Richtung eines Konstruktionslexikons. Anschliessend werden Analysen einzelner Konstruktionen aus den Korpora durchgeführt und die Schritte beschrieben, die zur Erfassung neuer und zur Ergänzung bestehender Lexikoneinträge führen.

Für den folgenden praktischen Durchführungsteil sollen noch einmal die praktischen Ziele und Schritte der Lexikonerstellung kurz zusammengefasst werden:

1. *Lexikoneinträge* sollen mit so vielen Informationen wie möglich versehen werden, die für die Erzähl- und Gesprächsanalyse nötig und nützlich sind. Diese Einträge haben prototypischen Charakter.
2. *Prototypische Beispiele* sollen aus den Korpora erschlossen, passende Bedeutungen festgelegt und im Lexikon abgebildet werden. Mit Hilfe der Lexikonmerkmale sollte ein erneutes Vorkommen dieser Form-Bedeutungseinheit korrekt im Text erkannt und allenfalls kodiert werden können.
3. Der erzählanalytische Hintergrund bringt eine Fokussierung auf die *Verben* mit sich (das Lexikon ist in diesem Sinne nicht ausgewogen).
4. Das Hauptgewicht soll auf Konstruktionen liegen, die in *Gesprächen* angewendet wurden. Jeder Eintrag soll aus konstruktionsgrammatischer Sicht motiviert sein. Dabei geht es um die Prüfung der Konzepte und deren Nützlichkeit auf Lexikonebene. Die Implementierung in der Analyseapplikation spielt erst in zweiter Linie eine Rolle.

5. Der Ausbau des Lexikons wird nicht nur theoretisch beschrieben, sondern gleichzeitig auch im online zugänglichen Lexikon implementiert.¹

Es folgt eine Beschreibung der neuen Merkmale und Funktionen, die im JAKOB-Lexikon implementiert wurden, um die in den bisherigen Kapiteln angesprochenen und teilweise bereits gemachten Ergänzungen und Neuerungen umzusetzen (Luder, Clematide & Distl, 2008; Luder, 2010; Luder & Clematide, 2010). In einer ersten Phase (2007/2008) wurde das bisherige JAKOB-Lexikon im OLIF-Standard (Kapitel 9.1.2) neu implementiert. Damit wurde – als wichtigste Neuerung – der Übergang von einem Einwort-Lexikon zu einem Mehrwort-Lexikon gemacht. In einer zweiten Phase wurde die OLIF-Implementierung um weitere Eigenschaften ergänzt, um den Anforderungen der Darstellung von Konstruktionen und der integrierten Verwendung von Verbpatterns (erweiterte Semantik) zu entsprechen. Anschliessend werden die Analysebeispiele mit den daraus entwickelten Lexikoneinträgen vorgestellt.

9.1 Neue Anforderungen an das JAKOB-Lexikon

In der Planungsphase, während der Implementierung und schliesslich in der Aufbau- und Anwendungsphase wurden die folgenden Leitsätze für das JAKOB-Lexikon entwickelt und den Anforderungen der Implementierung und der Erzählanalyse angepasst.

- Das JAKOB-Lexikon ist ein Lexikon für *Konstruktionen* und *Phraseologismen*, aber ebenfalls für *Einzelworte*, also nicht ausschliesslich ein Phraseologismen-Lexikon wie z. B. Phraseo-Lex (Keil, 1997).
- Lexikoneinträge sind *Form-Bedeutungspaare* im Sinne der Konstruktionsgrammatik (siehe Kap. 6.4).
- Lexikoneinträge sind *Prototypen*. Die *kanonische Form* (Nennform, Stammform, Lemma) des Lexikoneintrags ist nicht vollständig fixiert, sondern Variationen und Ergänzungen sind möglich, ohne dass dafür in der Stammform explizit Platzhalter gesetzt werden müssten (Wortreihenfolge, optionale Ergänzungen und Erweiterungen). Umgekehrt werden variable, aber obligatorische und für das Verständnis des Ausdrucks notwendige Ergänzungen in eckige Klammern gesetzt.

¹ siehe URL: <http://www.jakob.uzh.ch/lexikon/>

- *Valenz* und Nennform: Die Nennform muss *nicht* alle Valenzstellen beinhalten, die Nennform heisst z. B. «hinters Licht führen» und nicht «jmdn hinter Licht führen». Keil (1997, S. 63) kritisiert allerdings diese Schreibweise mit der Begründung, dass das obligatorische Akkusativobjekt nicht weggelassen werden dürfte. Die obligatorischen und teilweise die fakultativen Ergänzungen sind aber im JAKOB-Lexikon dreifach abgebildet, und zwar in den Merkmalen <synFrame>, <pattern> und <bauplan>, erscheinen aber nicht in der Nennform.
- *Schematische Konstruktionen*: Ins Lexikon können nur lexikalisierte oder mindestens teillexikalisierte Konstruktionen aufgenommen werden. Die gänzlich schematischen Konstruktionen (ohne lexikalisierte Komponenten, z. B. die *Transitivkonstruktion*) sind für die Verben in den Satzmustern nach Wahrig enthalten.
- Das *Verbpattern* (<pattern>) ist ebenfalls prototypisch und kann mit Ergänzungen versehen werden. Im Pattern vorhandene Elemente müssen aber realisiert sein, wenn sie nicht als optional bezeichnet sind.
- Lexikoneinträge sind aus dem «normalen Gebrauch» in den zur Verfügung stehenden Korpora abgeleitet (vgl. Hanks & Pustejovsky, 2005).

9.1.1 Lexikonaufbau: Lesarten

Am Beispiel des Funktionsverbs «leisten» sollen kurz verschiedene Möglichkeiten des Aufbaus von Lexikoneinträgen aufgezeigt werden. Das Beispiel ist Hanks et al. (2006) entnommen und zeigt auf Grund von Korpusevidenzen die Kookkurrenzen von «leisten» auf, die in vier Gruppen eingeteilt werden (ausführliche Darstellung im Kapitel 4.6.2):

1. *leisten mit direktem Objekt* (Bsp. «einen Beitrag leisten»).
2. *Reflexive Verwendung*: «sich etwas leisten».
3. *leisten mit abstraktem Objekt* («Der Motor leistet 200 PS»).
4. *leisten mit einem indefiniten Pronomen oder einem nominalisierten Adjektiv als direktem Objekt* («nichts Vergleichbares leisten»)

Die erste Möglichkeit besteht darin, dass die vier erwähnten Varianten des Funktionsverbgefüges als vier verschiedene Einträge realisiert werden, wie bei Hanks et al. (2006) vorgeschlagen, getreu dem Prinzip der Lexikoneinträge als *Form-Bedeutungspaare*. Ein praktisches Problem entsteht

dadurch für die Anzeige der Nennform im Lexikon ($\langle \text{canForm} \rangle$): Wie werden Ausdrücke dargestellt, die nicht gänzlich fixiert sind, sondern variable Anteile haben? Für diesen Fall (den Regelfall) werden die neuen Felder $\langle \text{pattern} \rangle$ und $\langle \text{bauplan} \rangle$ zusätzlich zur Nennform herangezogen, um die Entsprechung von Form und Bedeutung zu bestimmen.

Die zweite Möglichkeit würde darin bestehen, alle Lesarten von «leisten» unter der einfachen Nennform «leisten» aufzuführen (nach dem Prinzip des *sense enumerative lexicon*); in diesem Fall müsste der Lexikoneintrag intern in mehrere Unterbedeutungen (mit verschiedenen syntaktischen und semantischen Satzmustern) aufgespalten werden, was in gedruckten Wörterbüchern üblich ist, aber dem geplanten Lexikonaufbau widerspricht. Diese Darstellungsart wird für das vorliegende Projekt nicht weiter verfolgt.

Die Lesart 1) liefert bereits eine grosse Zahl verschiedener Bedeutungen, die vom (satzsemantischen) Standpunkt aus nicht in der gleichen Form-Bedeutungseinheit dargestellt werden können (Beispiel: Können «Hilfe leisten» und «Widerstand leisten» unter der gleichen Nennform subsumiert werden?)

Der *Lexikoneintrag* für «leisten einen Beitrag» könnte gestaltet werden, wie im Beispiel (40) ersichtlich. Die wichtigste Implementierungsfrage ist dabei die Funktion und die Form von *canForm*.

- (40) *canForm*: *leisten einen [MOD] Beitrag*
 Pattern: [[Human]] *leisten einen [MOD] Beitrag*
 Konstruktion: [einen ([ADV]) *Beitrag leisten*]
 «ich leiste einen bescheidenen Beitrag»

Ein Beispiel wie «lesen» zeigt aber, dass das in der Lesart 1) oben angesprochene Problem nicht nur für komplexere Konstruktionen wie die erwähnten Funktionsverbgefüge besteht, sondern auch schon für einfache Verben, die eine transitive und eine intransitive Lesart haben. Werden die beiden Lesarten explizit als zwei verschiedene Lesarten im Lexikon gespeichert oder wird die Akkusativvergrößerung als fakultativ bezeichnet? Um einer Konstruktionsinflation vorzubeugen, ist eine pragmatische Lösung angebracht, die fragt, wozu die Unterscheidung nötig ist.

9.1.2 OLIF Eigenschaften

Die Tabelle (9.1) zeigt als Beispiel einen Lexikoneintrag mit den wichtigsten Attributen der OLIF-Struktur, wie sie bei der Neuimplementierung des Lexikons verwendet wurden. Eine ausführliche Beschreibung aller OLIF-Merkmale findet sich bei McCormick (2005). Erläuterungen zu einzelnen

Merkmale folgen nachstehend und berücksichtigen hauptsächlich die Phänomene von linguistischem, psychologischem und lexikographischem Interesse, die technischen Details werden weggelassen.

Das Merkmal *JAKOB-Code* ist ebenfalls aufgeführt, obschon dieses natürlich nicht zu den OLIF-Eigenschaften gehört. Die Kodierung der Lexikoneinträge gemäss Kodiersystem ist eine lexikographische Aufgabe, die Vertrautheit mit dem JAKOB-Kodiersystem voraussetzt. Im vorliegenden Beispiel wird die Konstruktion <ein Zeug machen> mit einer Kombination der Codes DAR (darstellen) und ERR (erregen) kodiert. Die Konstruktion beinhaltet sowohl Aspekte des Präsentierens und der Darstellung (DAR) als auch solche der aggressiven und aversiven Stimulation (ERR).² Die Beschreibung der Verbcodes in den Konstruktionstabellen beinhaltet: Kürzel, Bezeichnung des Verbcodes, Dimension, Subdimension.

Tabelle 9.1: Beispiel OLIF: Lexikoneintrag

OLIF-Attribut	Wert	Beschreibung
canForm	<i>machen ein Zeug</i>	canonical form, Nennform, Lemma
orthVariant	machen ein Zeugs	orthografische Variante
crossRef	z. B. machen ein Theater	Verknüpfung (cross reference)
linkType	near-synonym	Typ der Verknüpfung
head	machen	head (Kopf) des Ausdrucks
ptOfSpeech	Verb	Wortart des Heads
entryFormation	phrase	Typ des Eintrags (phrase = MWE)
phraseType	idiom	Typ des MWE
synFrame	500	Satzmuster nach Wahrig: Verb + Akkusativobjekt
synType	function verb	Funktionsverb
semType	unspecified activity	semantische Typen OLIF
definition	sich kompliziert verhalten, Aufregung produzieren	
subjField	Therapiesgespräch	Diskurstyp (Domain, Genre)
JAKOB-Code	DAR-ERR	DAR=Darstellen (Handeln – Präsentieren); ERR=Erregen (Geschehen – Vorgehen)

Basic properties (Grundmerkmale)

Die *Nennform* wird in der OLIF-Terminologie <canForm> genannt und kann aus einem oder mehreren Wörtern bestehen. Der <head> (Kopf) eines Eintrags bezeichnet das syntaktische Zentrum des Ausdrucks und wur-

²Die Definition der einzelnen Verbcodes kann im JAKOB-Lexikon online durch Anklicken der Codelinks abgerufen werden.

de für verbale Ausdrücke auf das Verb festgelegt.³ Ausserdem steht per Konvention in der `<canForm>` der Kopf am Anfang des Ausdrucks (siehe Beispiel). Unterschieden wird zwischen Einzelwörtern und Mehrwortausdrücken (`<entryFormation>`: `<sgl>` vs. `<phr>`); für Phraseologismen können verschiedene Typen spezifiziert werden (u.a. `<multi word>`, `<set-phrase>`, `<idiom>`, `<metaphor>`).

OLIF erlaubt die Erfassung *einer* orthographischen Variante der Nennform innerhalb des gleichen Eintrags im Merkmal `<orthVariant>`. Eine freie Beschreibung der Konstruktionsbedeutung befindet sich im Merkmal `<definition>`.

Cross Reference

Die Eigenschaft `<cross reference>` ist ein mächtiges Merkmal, um Lexikoneinträge mit verschiedenen Relationstypen zu verknüpfen. Die Art der Verknüpfung zwischen dem Starteintrag und dem Zieleintrag wird durch die Eigenschaft `<linkType>` näher bestimmt. In der Tabelle (9.2) wird die Liste der `<linkType>`-Eigenschaften aus der OLIF-Dokumentation (McCormick, 2005, S. 31) ausschnittsweise zitiert, um das theoretische Potenzial der assoziativen Verknüpfungen, die von einzelnen Lexikoneinträgen ausgehen können, zu demonstrieren. Der praktische Nutzen wird allerdings relativiert, einerseits durch den grossen lexikographischen Aufwand, der für die Verlinkung nötig ist (alle Verknüpfungen müssen manuell erstellt werden), und andererseits dadurch, dass die Auswahl eines allgemein akzeptierten und nützlichen Verlinkungskonzeptes eine schwierig zu lösende Aufgabe ist. So wurden bis jetzt nur die wichtigsten und naheliegendsten Linktypen verwendet (z. B. `synonym`, `near-synonym`, `antonym`, `near-antonym`, `is-associated`).

In der OLIF-Dokumentation werden diese verschiedenen Verknüpfungstypen nicht näher umschrieben, was die Zuordnung weiter erschwert. So sind schon die Begriffe *synonym*, *near-synonym*⁴ und *antonym*, deren unterschiedliche Bedeutungen intuitiv einleuchten, schwierig zu definieren, wie das folgende Zitat zum Aspekt des *lexikalischen Kontrastes* zeigt:

³im Gegensatz zu Lexikonkonzepten, die z. B. in nominalen Funktionsverbgefügen das Nomen als Kopf bezeichnen.

⁴«Plesionymie wird häufig auch unter dem Begriff Beinahe-Synonymie (engl. near-synonymy) aufgeführt (Cruse, 2004). Der Austausch eines Wortes durch sein Plesionym hat Auswirkung auf den Wahrheitswert einer Aussage. Das Paar *foggy* – *misty* sind beispielsweise Plesionyme [...]. Man kann zugleich das eine assertieren und das andere verneinen, ohne dass dies zu einer paradoxen Aussage führt.» (Roth, 2006, S. 9)

Tabelle 9.2: <linkType>-Werte für <cross reference> (OLIF, Ausschnitt)

linkType	Beschreibung
synonym	synonym of
near-synonym	near synonym of
antonym	antonym of
near-antonym	near antonym of
has-hyperonym	is kind of (subordinate) (Pferd has hyperonym Säugetier)
has-hyponym	is kind of (superordinate)
has-holonym	part of
has-meronym	whole of
causes	cause of
is-caused-by	effect of
has-subevent	(between verbs/gerunds) e.g., sleep > snore
is-subevent-of	(between verbs/gerunds) e.g., snore > sleep
role	activity that something (noun) is involved in
involved	thing (noun) involved in activity represented by verb
produces	producer of
is-product-of	product of
is-spatial-rel	related spatially
is-associated	associated term
is-child-of	offspring of
is-parent-of	parent of
manner-of	(adv > verb) noisily > snore
no-synonym	not allowed as synonym
has-no-syn	has disallowed synonym
fuzzynym	(noun > noun; verb > verb) fuzzy semantic relation
un	relation unspecified

Lexical contrast implies lexical similarity. Antonyms, it has been said, are word pairs that are semantically <minimally different> (Clark, 1970; Hale, 1971; Murphy, 2003). Given such a definition, it would seem that nothing would make better antonyms than near-synonyms, which are more similar than they are different. Indeed, Petra Storjohann's «Plesionymy: a case of synonymy or contrast?» (Storjohann, 2009) makes the case that it is as usual for near-synonyms to be used in contrast with one another as to be used as equivalents. The profusion of near-synonymy and the avoidance of true synonymy in language (languages abhor absolute synonyms just as nature abhors a vacuum, Cruse, 1986:270) support the view that contrast is central to the pragmatic process of meaning construal (e.g.

Breal, 1900; Clark, 1992). Storjohann furthermore shows that plesionyms are construed as having contrasting meanings when they co-occur in the types of lexico-grammatical frames in which canonical antonyms typically co-occur (Jones, 2002), thus supporting Murphy's (2006) argument that these frames are meaningful constructions, i.e. contrastive constructions. (Murphy, Paradis & Willners, 2009, S. 2137)

Für die praktische Zuweisung dieser drei Verknüpfungstypen heisst dies bis jetzt, dass Antonyme wenig, Synonyme kaum vorkommen, und dass Beinahe-Synonyme eine Art Restkategorie zwischen den beiden erstgenannten bilden. Eine wichtige Aufgabe für die weitere Entwicklung des JAKOB-Lexikons wird es sein, die Rolle der Verknüpfungen (cross references) im Lexikon in zweierlei Hinsicht genauer festzulegen. Der erste Punkt betrifft die Anzahl der Verknüpfungen; schon bei einer sparsamen Zuordnung ergeben sich mehrere Verknüpfungen pro Eintrag. Wie viele davon sind notwendig? Zum zweiten stellt sich die bis jetzt noch nicht beantwortete Frage, wie die Verknüpfungen bei der Suche im Lexikon angezeigt und bei der Textanalyse eingesetzt werden sollen.

Syntaktische Merkmale

Zu den syntaktischen Merkmalen gehören die *Wortart* <ptOfSpeech> und der *syntaktische Typ* <synType> (für Verben: Hauptverb, Modalverb, Funktionsverb, Hilfsverb). Die *Satzmuster* für Verben <synFrame> wurden – wie im bisherigen JAKOB-Lexikon – aus dem «Kleinen Wahrig» (Wahrig, 2007) übernommen, da diese eine genauere Kategorisierung als die OLIF-Originalwerte erlauben. Die Satzmuster nach Wahrig beschreiben 76 verschiedene Valenzpattern für Verben mit Hilfe einer dreistelligen Zahl; im *Kleinen Wahrig* sind den Verben und ihren Lesarten die entsprechenden Muster zugeordnet (siehe Beispiele 41).

- (41) a. 400: VERB – KEIN OBJEKT («schlafen»)
 b. 500: VERB + AKKO (transitive Verben)
 c. 505: VERB + AKKO + (PRÄPO) (Akkusativobjekt und fakultatIVES Präpositionalobjekt, z. B. «befragen»)
 d. 600: VERB + DATO («helfen»)

Semantische und pragmatische Merkmale

OLIF stellt eine Liste von *semantischen Typen* <semType> zur Verfügung, mit denen Einträge charakterisiert werden können, und auch eine grössere Liste zur Kennzeichnung der *Wissensdomänen* <subjField>; beide Wertelisten stammen aus den Bereichen Wirtschaft und Industrie und sind deshalb nur bedingt für das vorliegende Projekt brauchbar. Tabelle (9.3) listet die semantischen Typen für Verben auf (McCormick, 2005).

Tabelle 9.3: Semantische Verbklassifikation in OLIF

achiev	achievement
act	unspecified activity
emot	emotion
event	event
ment-act	mental activity
mov	movement
– mov_motdir	directed motion, e.g., dance, depart, fly
– mov_motnd	non-directed motion, e.g., depart, go, walk
noise	noise-producing
phys-act	physical activity, e.g., persist, refrain, appear
percept	perceptive
perm	permission verb
pha	phrasal verb
pro	process
sense	sense
situat	situation
stat	stative, e.g., grow, become, sound

Weil semantische und vor allem pragmatische Eigenschaften von Lexikoneinträgen nur eingeschränkt kontextfrei bestimmt werden können, entsteht bei diesen Eigenschaften die grösste Notwendigkeit, nach neuen Lösungen für das veranschlagte Projektziel – die Repräsentation von Konstruktionen als Form-Bedeutungspaare – zu suchen. So ist z. B. der semantische Typ eines Verbs zwar bis zu einem gewissen Grade aufschlussreich für eine Typologisierung des Verbes selbst, wichtiger für den Satzzusammenhang sind aber die semantischen Typen der Argumente, z. B. der Nomen, die mit diesem Verb zusammen auftreten. Ontologien beziehen sich hauptsächlich auf Nomen; die Tabelle (9.4) zeigt die semantischen Typen, die OLIF für *Nomen* bereitstellt:

Die Werte, die OLIF für die pragmatische Eigenschaft <subjField> bereitstellt, wurden bereits auszugsweise in der Tabelle (7.5, Seite 175) vorgestellt; ausführlich sind sie bei McCormick (2005) beschrieben.

Tabelle 9.4: Semantische Nomenklassifikation in OLIF

abs	abstract, e.g., format, rapidity, poverty, type
– abs-ag	abstract agent, e.g., efficiency, cause, method, goal, event
– abs-gen	general abstract concept, e.g., truth, idea, justice
– abs-nonag	non-verbal abstract, e.g., shape, condition, class, feature
anim	animate, e.g., manager, committee, subscriber, buyer
– anim-ani	animal, e.g., deer, bacteria, gnat, weasel
– anim-hum	human, e.g., employee, scientist, Professor, Mrs.
– – anim-hum-func	office, title, e.g., Dr., President, General
– – anim-hum-pn	human proper name, e.g., John, Mr. Smith, Marie
– anim-soc	social institution, e.g., agency, company, bureau, business
– – anim-soc-org	specific organization, e.g., EC, United Nations, NASA
asp	aspective, e.g., prototype, majority, piece
cnc	concrete, e.g., table, battery, ligament, missile
– cnc-ag	concrete agent, e.g., camera, radio, truck, explosives
– cnc-amor	amorphous, e.g., breeze, tide, atmosphere
– cnc-atom	atomistic, e.g., electron, granule, nucleus
– cnc-class	classifier, e.g., compound, substance, element
– cnc-color	color, e.g., olive, orange, cherry
– cnc-ednm	edible (non-mass), e.g., cracker, lemon, pork chop
– cnc-func	functional, e.g., box, wall, pipe, circuit, shirt
– cnc-light	impulse/light, e.g., beacon, ray, tone, flare
– cnc-mark	mark/blemish, e.g., boil, blemish, scratch
– cnc-nat	natural, e.g., cloud, pebble, flower
– – cnc-nat-plant	plant, e.g., violet, clove, lilac
inform	information, e.g., newspaper, symbol, rule, ballistics
– inform-sen	semiotic system, e.g., address, signal, code, number
loc	locative, e.g., office, zone, city, room, Munich
mass	mass, e.g., iron, water, sand, fiber, fire, heat
– mass-mat	material, e.g., aluminum, wool, plastic, glass
meas	measure, e.g., pressure, quantity, gram, rpm, voltage
– meas-abs	abstract measure, e.g., temperature, length, velocity
– meas-disc	discrete measurable concept, e.g., increment, sum, count
– meas-unit	unit of measure, e.g., inch, cm, hour, volt, hertz, kph
proc	process, e.g., correction, analysis, call, removal
tmp	temporal, e.g., summer, morning, September, Friday

Nicht umgesetzte OLIF-Möglichkeiten

Einige Merkmale, die OLIF bietet, wurden in der vorliegenden Implementierung nicht umgesetzt, so sind zum Beispiel keine morphologischen Eigenschaften erfasst worden, weil die mit der Erzählanalyse JAKOB untersuchten transkribierten Texte vorgängig segmentiert, morphologisch analysiert

und getaggt und lemmatisiert werden. In der OLIF-Struktur sind morphologische Eigenschaften vorhanden, jedoch in der aktuellen Implementierung ausgeblendet. Das Merkmal `<morphStruct>` bildet z. B. die morphologische Struktur ab.

Neben der Möglichkeit, Kreuzverweise zu erstellen, beinhaltet OLIF auch eine *Transferkomponente*, mit der korrespondierende Ausdrücke aus verschiedenen Sprachen nach dem gleichen Prinzip wie die *cross references* verknüpft werden können (z. B. für Übersetzungssysteme). Diese Komponente wurde nicht in die Datenbank implementiert.

OLIF ermöglicht von der Datenstruktur her einerseits Lexikoneinträge, die Wörter oder Ausdrücke repräsentieren, darüber hinaus können Einträge aber auch als *Terme* oder *Konzepte* aufgefasst werden (Eigenschaft `<entryStatus>`), so dass implizit eine Ontologie quasi über das Lexikon gestülpt werden könnte. Im vorliegenden Projekt kommt nur der Wortstatus zur Verwendung.

9.1.3 Neue Merkmale im Lexikon

Wie in den vorausgegangenen Kapiteln erwähnt, wurde die OLIF-Struktur im Laufe des Aufbaus (2008-2010) mit weiteren Merkmalen ergänzt, die für die Repräsentation von Konstruktionen im vorgesehenen Kontext sinnvoll schienen. Einige der «neuen» OLIF-Eigenschaften haben das Lexikon einen grossen Schritt in Richtung auf eine lexikalische Ressource für Konstruktionen vorwärts gebracht, andere Ergänzungen haben (noch) nicht die erhofften Verbesserungen gebracht, und ihr zukünftiger Nutzen ist noch nicht klar.

Die Tabelle (9.5) zeigt die für das vorliegende Lexikonprojekt neu definierten Lexikoneigenschaften zusammenfassend, im anschliessenden Text werden sie genauer beschrieben. Die Merkmale `<natclause>`, `<bauplan>`, `<pattern>` und `<construction>` bezeichnen syntaktische und semantische Eigenschaften des Lexikoneintrags.

Natclause und Bauplan

Die Merkmale `<natclause>` und `<bauplan>` wurden eingeführt, um die syntaktische Struktur der Verbpatterns mit Hilfe eines Parsers bestimmen zu können. Das Merkmal *natclause* dient als Eingabe für den an der Universität Zürich entwickelten Parser *Pro3GresDe* (Schneider, 2008; Senrich,

Tabelle 9.5: Neue semantische Merkmale im Lexikon, Bsp. «lesen»

Attribut	Wert	Beschreibung
(canForm)	<i>lesen</i>	
natclause	«jmd liest etwas» (Erläuterung im Text)	prototypisches natürlich-sprachliches Satzmuster
bauplan	500: 1 - - L2 subj Human 2 lesen VVFIN L0 root 3 - - L2 obja Anything	maschinenlesbares Satzmuster
pattern	[[Human]] lesen [[Document]]	Verbpattern (nach Hanks, 2008a)
construction	Konstruktion	

Schneider, Volk & Warin, 2009).⁵ Die Ausgabe des Parsers wird in das Feld «bauplan» geschrieben, kann dort weiter bearbeitet werden und dient als Vorlage für die manuelle Bearbeitung des Merkmals «pattern».

Als Eingabe für das Feld «natclause» formuliert die Lexikographin einen einfachen Satz, der syntaktisch das in Bearbeitung befindliche Verb repräsentiert und die Valenzstellen abbildet. Dabei werden für die variablen Valenzstellen vordefinierte allgemein gehaltene Platzhalter eingesetzt, in der aktuellen Testphase beschränkt auf die folgenden Marker für Infinitpronomen in verschiedenen Kasus: «jmd», «jmdn», «jmdm», «etwas», «etwasn», «etwasm».⁶ Die Platzhalter werden nach dem Parsing durch semantische Typen ersetzt. Das Beispiel (42) zeigt die Eingabe für das transitive Verb «lesen» (im Eingabefeld «natclause») mit den Platzhaltern für Subjekt und Akkusativobjekt und in der zweiten Zeile die *Parseerausgabe* (in das Feld «bauplan») im Textformat. Die Abbildung (9.1) zeigt die alternative grafische Ausgabe des Parsers zum besseren Verständnis als Baumdarstellung (ohne Umwandlung der Platzhalter).

(42) *jmd* *liest* *etwasn*
 500: 1 - - L2 subj Human 2 lesen VVFIN L0 root 3 - - L2 obja Anything

Das Beispiel (43) zeigt die Eingabe/Ausgabe für den idiomatischen Ausdruck «an den Karren fahren». Für Verben mit abtrennbarem Präfix wird die Eingabe als «weil»-Satz gestaltet, damit der Parser Verb und Präfix korrekt kombiniert, wie im Beispiel «Dampf ablassen» (44) gezeigt. Die

⁵Pro3GresDe – a Dependency Parser for German.

URL: <https://files.ifi.uzh.ch/cl/gschneid/parser/>

⁶Der Parser wurde zu diesem Zweck entsprechend angepasst.

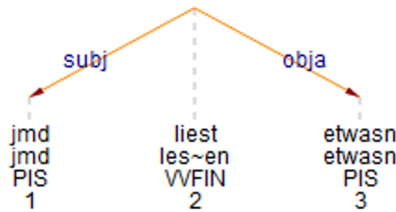


Abbildung 9.1: Baumdarstellung der Parserausgabe für das Verb «lesen»

Konjunktion «weil» wird bei der Verarbeitung des Rückgabestrings weggelassen.

- (43) *jmd fährt jmdm*
 605: 1 - - L2 subj Human 2 fahren VVFIN L0 root n/a 3 - - L2 objd Human
an den Karren
 4 an APPR L2 pp n/a 5 die ART L6 det n/a 6 Karren NN L4 pn

- (44) *weil jmd Dampf ablässt*
 500: 1 - - L3 subj Human 2 Dampf NN L3 obja n/a 3 ablassen VVFIN L0 root

Der Rückgabestring des Parsers wird folgendermassen aufbereitet und in das Feld «bauplan» geschrieben (siehe Beispiele 42 bis 44)⁷: Die Platzhalter «jmd» und «etwas» werden durch zwei Werte aus der CPA-Ontologie ersetzt (die Ontologie, die für den Aufbau des Feldes «pattern» verwendet wird), und zwar wird «jmd» (Kasus beliebig) durch «Human» ersetzt, «etwas» durch «Anything». Die Aufgabe des Lexikographen ist es dann, «Anything» durch einen geeigneten semantischen Typ aus der CPA-Ontologie zu ersetzen (z. B. «Human», «Institution»), passend zur Semantik des Verbausdrucks. Natürlich könnte die Komplexität des Matching beliebig ausgebaut werden. Am Anfang des Rückgabestrings wird das Satzmuster gemäss Wahrig vorangestellt. Die Zeichenkette «bauplan» besteht also aus der dreistel-

⁷Dieser Vorgang wird interaktiv während der Bearbeitung des Datensatzes durch Anklicken eines Buttons ausgelöst.

ligen Zahl für das Satzmuster (syntaktische Valenz), gefolgt von 6 Werten für jedes geparste Wort:

1. *Laufnummer* (Wortnummer), Position dieses Wortes im String.
2. *Grundform des Wortes*. Im Falle der Platzhalter «jmd» und «etwas» wird das Wort unterdrückt (weil variabel) und durch Bindestrich ersetzt; diese Position wird stattdessen durch den semantischen Typ (Pos 6) definiert.
3. *Wortart* (gemäß STTS-Tagset).⁸ Auch die Wortart wird bei den Platzhaltern unterdrückt und durch Bindestrich ersetzt.
4. *Link* zur Laufnummer des Root-Elementes (vgl. Baumdarstellung des Beispiels «lesen»). Im Sinne der Dependenzgrammatik ist das Verb die Wurzel der Satzstruktur; Subjekt und Akkusativobjekt besitzen deshalb einen Link zur Position 2 (root).
5. *Syntaktische Informationen* (Konstituente/Satzglied, Kasus, Valenzstelle), z. B. subj=Subjekt, obja=Akkusativobjekt. Die Satzglieder in der Satzstruktur müssen mit dem vorangestellten Satzmuster übereinstimmen.
6. *Semantischer Typ* gemäß CPA-Ontologie. Wichtig für die Erzählanalyse ist die Unterscheidung von menschlichen Akteuren («Human») gegenüber Dingen und Sachverhalten (auf der obersten Ontologieebene «Anything»), Kategorisierungsmöglichkeiten sind im Kapitel (7.3) beschrieben. Die Bestimmung des semantischen Typs für alle Elemente eines Verbpatterns ermöglicht die Unterscheidung von Bedeutungsvarianten in ähnlichen Verbpattern.

Die Eigenschaft «bauplan», entwickelt nach dem Vorbild der Darstellungsf formalismen im niederländischen Lexikon für Mehrwortausdrücke (DuELME, siehe 7.2.7) enthält damit alle Konstruktionskomponenten des Lexikoneintrages in maschinenlesbarer Form; syntaktische und semantische Eigenschaften sind hier kombiniert. In der Regel ist nach der automatischen Generierung des Bauplans eine manuelle *Nachbearbeitung* nötig. Die Nachbearbeitung umfasst vor allem die Präzisierung der semantischen Typen, sofern dies für das betreffende Muster adäquat ist. In diesem Sinne wird im

⁸Das Stuttgart-Tübingen Tagset (STTS) umfasst 54 Wortartenbezeichner (tags) für deutsche Textkorpora mit Wortartenannotation.

URL: <http://www.sfb441.uni-tuebingen.de/a5/codii/info-stts-de.xhtml> (15.11.2010)

Bauplan für «lesen» die Objektposition präzisiert und «Anything» durch «Document» ersetzt (aus der CPA-Ontologie, Kap. 7.3.3). Auch allfällige *syntaktische Restriktionen* müssen in diesem Schritt manuell ergänzt werden, wie z. B. in der Konstruktion «Angst haben vor», die (im «normalen» Gebrauch) Angst im Singular voraussetzt (Bauplan Bsp. 45).

- (45) 550: 1 - - L5 subj Human 2 Angst[sg][uncountable] NN L5 obja n/a 3 vor APPR
L5 pp - 4 - - L3 pn Anything 5 haben VA L0 root n/a

Die Eigenschaft «bauplan» soll für NLP-Anwendungen alle erforderlichen Parameter bereitstellen, ist aber kompliziert und wenig intuitiv zu lesen und könnte aus diesem Grund auch vor den Lexikonbenutzern verborgen werden. Für das intuitive Verständnis der Verbpatterns besser geeignet ist das Merkmal «pattern», das im Folgenden beschrieben wird.

Pattern (Verbpattern)

Im Verbpattern nach dem Vorbild der CPA (Corpus Pattern Analysis, Hanks, 2008a) sind die semantischen Eigenschaften der Patternelemente und deren Zusammenspiel abgebildet. Verbpatterns repräsentieren die semantischen Eigenschaften aller Elemente einer (verbalen) Konstruktion; die Bedeutung des Patterns erschliesst sich aus den semantischen Typen der einzelnen Elemente. Damit ergibt sich eine wesentlich genauere Beschreibung des Ausdrucks als allein durch die OLIF-Eigenschaft «semType», die den semantischen Typ des Ausdrucks als Ganzes repräsentiert. Die folgenden Beispiele (46-48) zeigen die Verbpatterns für die oben erwähnten Beispiele (42-44).

- (46) *jmd liest etwasn (canForm: lesen)*
[[Human]] lesen [[Document]]
- (47) *jmd fährt jmdm an den Karren (canForm: fahren an den Karren)*
[[Human]] fahren [[Human]] {an den Karren}
- (48) *weil jmd Dampf ablässt (canForm: ablassen Dampf)*
[[Human]] ablassen {Dampf}

Wichtige Eigenschaften der Verbpattern für die Herstellung von Bedeutung sind:

- Die semantischen Typen der Patternelemente sind prototypische Füller (filler/slots), und nicht als obligatorische Restriktionen anzusehen.
- Die prototypischen Kategorien und ihre semantischen Kernbedeutungen können im kreativen Sprachgebrauch durch «Sinnerweiterungen» (exploitations) überschrieben werden.
- Zusätzliche Ergänzungen und Angaben können das Pattern und zum Teil auch die Patternelemente erweitern.

Die Unterscheidung des *normalen* vom *kreativen* Sprachgebrauch ist theoretisch einleuchtend, praktisch jedoch oft schwierig zu beurteilen. So kann «ein Buch lesen» als normaler Sprachgebrauch bezeichnet werden; auch «Gedanken lesen» gehört als geprägter Ausdruck dazu. Schon schwieriger wird die Beurteilung bei Ausdrücken wie «viel Lob lesen» und «viel Kritik lesen» (Beispiele im deWAC). «Augen lesen» kann als *exploitation* bezeichnet werden (Beispiel 49). Die Verwendung von *metonymischen Ausdrücken* wie in den Beispielen (50) und (51), *einen Punkt lesen* oder *den Spieler lesen* zeigt einerseits, wie wichtig der Kontext für das Verständnis ist (worauf bezieht sich der «Punkt»?) und andererseits führen diese beiden Beispiele vor Augen, wie variabel die sprachlichen Ausdrucksmöglichkeiten durch die Verwendung von Metonymien werden.

- (49) P: so also, mich dünkt auch ich könne helle Augen besser lesen als dunkle [...] ich meinte es hange irgendwo auch an dem... (Korpus WIL).
- (50) P: ich möchte da wissen was darunter zu verstehen ist oder es hat einen Punkt drin äh : den man eben auch lesen kann wie man will... (Korpus WIL).
- (51) versuchte ich äh den den Spieler von Dostojewski zu lesen... (Korpus WIL).

Eine für die praktischen Lexikographiebelange brauchbare Beurteilungsgrundlage ist die Häufigkeit des Vorkommens der zur Diskussion stehenden Wortverbindungen in passenden Korpora.

Construction

Die Vorgabe für die Entwicklung dieses Merkmals lautete ursprünglich: Eine Darstellung des Lexikoneintrages als Konstruktion, die nicht maschinenlesbar sein muss, die als Konstruktion von Konstruktionsgrammatikern

gelesen werden kann und die alle Eigenschaften der Konstruktion, wie sie im Eintrag vorkommen, als String abbildet. Die letzte Vorgabe ist bereits im oben erwähnten *⟨bauplan⟩* realisiert und ergibt eine sehr kompliziert zu lesende Struktur. Es gibt keinen einheitlichen Darstellungsformalismus für Konstruktionen; die in der Literatur beschriebenen Beispiele lassen sich nicht auf einen gemeinsamen Nenner bringen (siehe 6.9). Das aktuelle Lexikonkonzept lässt sich am ehesten so beschreiben, dass *alle Merkmale* des Lexikoneintrags gemeinsam die Konstruktion charakterisieren. Die oben erwähnten neuen Eigenschaften sind dabei wichtige Konstruktionseigenschaften.

Die Idee, das Lemma (*⟨canForm⟩*) des Lexikoneintrags als Konstruktion zu bezeichnen und entsprechend zu formulieren, wurde verworfen, weil damit die Lesbarkeit der Nennform verschlechtert würde, so müssten z. B. Konstruktionen auf Satzebene auch die Subjektposition beinhalten. Das Merkmal Konstruktion wird also einstweilen als String zwischen eckigen Klammern formuliert, der syntaktische Informationen über die darin nicht explizit enthaltenen Argumente beinhaltet, wie in den Beispielen (52). Für Lexikoneinträge (Verben), die einem abstrakten Konstruktionsschema angehören, wird dieses abstrakte Schema als Konstruktion angezeigt.

- (52) a. *«lesen»*: Transitivkonstruktion [NP_{nom} V_{fin} NP_{akk}]
 b. *«Nägel mit Köpfen machen»*: [[NP_{nom}] machen Nägel mit Köpfen]

Wichtig als Ergänzung zur eher syntaxlastigen Konstruktionsbeschreibung im Merkmal *⟨construction⟩* ist hier das *Verbpattern*, das eine weniger weit gehende Schematisierung zulässt und sich auf die semantischen Aspekte konzentriert. Letztlich ist es sinnvoll, den ganze Lexikoneintrag als Konstruktion zu betrachten, da er wesentlich mehr Informationen beinhaltet als das Merkmal *⟨construction⟩* allein. Deshalb wird im Folgenden auch nicht ein Merkmal *⟨construction⟩* angestrebt, dass alle Konstruktionsaspekte beinhaltet.

Pragmatische Eigenschaften: SubjectField, Texttyp, Register, Topic

Die Tabelle (9.6) zeigt die neuen Merkmale *⟨subjField⟩*, *⟨textType⟩*, *⟨register⟩* und *⟨topic⟩*, die pragmatische und Diskurseigenschaften repräsentieren (*«Form, meaning and function of a text/discourse»*, Östman, 2005).

Eine ausführliche Auseinandersetzung mit verschiedenen pragmatischen Markern und entsprechenden Typologien wurde im Kapitel (7.3) vorge-

Tabelle 9.6: Neue pragmatische Merkmale im Lexikon

Attribut	Wert	Beschreibung
subjField	general	Domain, Genre (Gesprächsebene): Gesprächssorte oder Textsorte.
textType	description	Typisierung kleinerer Gesprächsabschnitte, auch als Diskurstyp oder Kommunikationstyp bezeichnet, z. B. Narrativ, Beschreibung oder Argumentation.
register	salopp, umg	auch als Stil (style) oder soziale Situation umschrieben, bezeichnet lokales Diskursmuster
topic	Stimmung, Sozialverhalten	beschreibt mit inhaltlichen Kategorien das Thema des lokalen Gesprächsabschnitts (z. B. Sport, Politik, Theater).

stellt. Aus den unzähligen Kategorisierungsvorschlägen und Begriffsdefinitionen wurden für das vorliegende Projekt *vier Kategorien* gebildet, um Eigenschaften des Kontextes, der Situation, der Kommunikationsart, des Gesprächstyps etc. abzubilden. Alle Äusserungen repräsentieren gleichzeitig verschiedene Beschreibungsebenen:

Multiple Determination: Every utterance (in speaking or writing) simultaneously exemplifies dialect, register, genre, and conversational variation. (Ferguson, 1994, S. 25)

Die vier Kategorien sollten sich vor allem in ihrer Reichweite unterscheiden (von der Charakterisierung kleiner Gesprächsabschnitte bis hin zum Gesprächs- und Diskurstyp).⁹

Eine prägnante *Abgrenzung der Konzepte* ist schwierig zu bewerkstelligen. Auch die Zuordnung der Kategorien zu den Lexikoneinträgen hat sich als ausserordentlich schwierig erwiesen, da sie hauptsächlich Eigenschaften beschreiben, die im weiteren und engeren *Kontext des Gesprächs* entstehen oder zur Wirkung kommen, während sie im (dekontextualisierten) Lexikoneintrag nicht eindeutig zu bestimmen sind. Es handelt sich deshalb um Lexikonmerkmale, die versuchsweise eingebaut und angewendet werden, deren Nutzen aber nicht von vornherein klar ist, ganz im Gegensatz zu den oben beschriebenen semantischen Eigenschaften.

⁹Die Kategorisierung erfolgt hier schwerpunktmässig für Gespräche und müsste für geschriebene Texte erweitert oder angepasst werden.

Subject Field (Domain, Genre)

Das Merkmal <subjField> gehört zu den originalen OLIF-Eigenschaften und wird hier noch einmal aufgeführt, um die Ergänzung der pragmatischen Eigenschaften eines Lexikoneintrages in der Zusammenschau zu sehen. Das Merkmal wird in der Literatur auch als *Genre* oder *Domain* bezeichnet und beschreibt mit extralinguistischen Kriterien den formalen Gesprächstyp/Diskurstyp, z. B. Tischgespräch, Märchen, Arztgespräch, Psychotherapiegespräch (Kuiper, 2009, S. 10). Man spricht hier auch von der *Makroebene* der Gesprächstypologie (Sager, 2001). Sager verwendet die Begriffe *Gesprächstyp* und *Gesprächssorte*. Östman (2005, S. 130) bezeichnet Rezepte, Todesanzeigen, Tischgespräche und Arztgespräch als Genres, die das *Setting* beschreiben («genres as contextual settings»). In der Auswahlliste des Archivs für Gesprochene Sprache gibt es weitere Gesprächstypen wie Beratung, Bericht, Dialog, Diskussion, Gespräch, Interview, Reportage, Schlichtung, Unterhaltung, Vortrag.

Aktuell im JAKOB-Lexikon zur Verfügung stehende Kategorien sind: *Therapiegespräch*, *formelles Gespräch*, *informelles Gespräch*, *Arztgespräch*, *Diskussion*, *Interview*, *Tagebuch* und *general* als unspezifizierte Gesprächsvariante.

Text Type (discourse type, communication type)

Der *Texttyp* beschreibt die kommunikativen Funktionen und die Typisierung kleinerer Gesprächsabschnitte. Verschiedene kommunikative Funktionen können nacheinander im gleichen Gespräch vorkommen. So können sich in einem Therapiegespräch narrative Episoden abwechseln mit Beschreibungen oder mit argumentativen Gesprächsstellen. Diese Texteinheiten werden durch interne linguistische Eigenschaften hergestellt, wie z. B. durch narrative, beschreibende oder argumentative Redeweisen (Östman, 2005). Texttypen können z. B. argumentativ, narrativ, instruktiv, expositorisch oder deskriptiv sein; dies gilt sowohl für geschriebenen Text als auch für gesprochenen Diskurs. Imo (2007) nennt diese Kategorie *kommunikative Praktiken*. Östman (2005) spricht von *Diskurspatterns* als konventionalisierte Assoziationen zwischen Text-Typ und Genre (S. 130).

Im JAKOB-Lexikon stehen aktuell die folgenden Werte zur Auswahl: *narrative*, *description*, *exposition*, *argumentation*, *instruction*.

Register (style, social situation, local patterns)

Als *Register* wird die Art der Gestaltung kleiner Diskurseinheiten (von lexikalischen Items bis zu Sätzen) bezeichnet; Sprecher geben ihren Äußerungen dadurch eine bestimmte Note für eine bestimmte Situation. Beispiele für Register-Eigenschaften sind: formell, informell, humoristisch, ironisch. Register wird auch als *Sprachstil* bezeichnet und gibt Hinweise auf die soziale Situation (Lee, 2001). Deppermann und Spranz-Fogasy (2001) verwenden die Bezeichnung *Modalität des Gesprächs* (z. B. Spass, Nacherzählung).¹⁰

Im JAKOB-Lexikon stehen die folgende Kategorien zur Verfügung: *formell, informell, umg (umgangssprachlich), salopp, derb, humoristisch, ironisch, gehoben, baby-talk*.¹¹

Topic, Thema

Die Eigenschaft *Topic* enthält inhaltliche Hinweise zum *Thema*, in dem eine Konstruktion verwendet wird. Geht es um einen Streit? Oder wird in der Küche gekocht? Oder wird eine Szene aus dem Büroalltag erzählt? Topic bezeichnet das *Thema* oder den *Kontext* der aktuellen (lokalen, kleinen) *Gesprächseinheit*, gibt ein Stichwort zur thematischen Situation (Bsp. Medizin, Kochen, Beruf, Sport). Topic soll es ermöglichen, verschiedene Bedeutungen eines Ausdrucks im Kontext mit *inhaltlichen Kriterien* zu unterscheiden. So können zum Beispiel die unterschiedlichen Bedeutungen des Ausdrucks «Ausschlag» im Beispiel (53) einerseits durch das Topic «Medizin» unterschieden werden (b), andererseits ist eine Disambiguierung auch möglich durch das Erkennen des Phraseologismus «den Ausschlag geben» (a).

- (53) a. Den Ausschlag gegeben hat dann die folgende Situation...
b. Auf der Haut zeigte sich ein fleckiger Ausschlag...

Das Merkmal Topic enthält sehr heterogene Werte, die Kriterien für die Liste sind noch nicht klar festgelegt. In der Testphase wurden viele Kategorien aufgenommen, um später eine Auswahl der nützlichen treffen zu können. Topic könnte teilweise auch mit den *Frames* (sensu FrameNet) in Verbindung gebracht werden.

¹⁰In der Erzählanalyse JAKOB ist der Begriff Modalität für die Beschreibung der *Segmenteigenschaften* reserviert. Segmente (Sätze) haben z.B. die Modalitäten *Fragesatz, Imperativ, Negation, modale Aussagen* (mit Modalverben) etc.

¹¹Da sich diese Merkmale im Versuchs- und Teststadium befinden, enthalten die Auswahllisten heterogene Werte und Konzepte, und sowohl deutsche als auch englische Ausdrücke.

Die folgende Liste zeigt eine Auswahl der für das Merkmal Topic aktuell verwendeten Begriffe:

arbeit, befinden, beruf, bewegung, bewusstsein, beziehung, einkaufen, essen, familie, ferien, freizeit, fremd, gefühle, geschäft, gesetz, gesundheit, kleidung, kochen, kunst, körper, liebe, macht, medizin, musik, reisen, religion, restaurant, schule, sex, sozial, spielzeug, sport, streit, tageszeit, technik, theater, therapie, tod, traum, verkehr, wetter

Weitere neue Merkmale

Das Merkmal *Frame* zeigt für den Lexikoneintrag den passenden Frame aus FrameNet oder aus der deutschen FrameNet-Variante SALSA¹² an. Dies kann allenfalls zu einem vertieften Verständnis und zu einer Reflexion über die Strukturen der Verbpatterns beitragen. FrameNet beruht allerdings auf einem ganz andern Hintergrund als die Verbpattern; eine Gegenüberstellung findet sich bei Pustejovsky et al. (2004, S. 2), siehe Kapitel (6.11.2).

Die Kategorisierung nach *Dornseiff* (Der deutsche Wortschatz nach Sachgruppen, Dornseiff, 2004) wird als Merkmal aufgenommen, weil der spezielle Aufbau dieses onomasiologischen Wörterbuches für den Aufbau der Lexikoneinträge oft gute Hinweise liefert.

Imo (2007) verwendet in seinen Konstruktionsbeschreibungen zusätzlich weitere Eigenschaften:

- *Funktion der Konstruktion*: beschreibt prototypische und kontextspezifische Funktionen der Konstruktion. So hat die teilspezifische Konstruktion <wenn ich (es) recht verstehe> die Funktion, «Äusserungen als nicht sicher zu markieren» (S. 281).
- Das Merkmal *Sequenz* liefert Informationen über den sequenziellen Rahmen, die Gattung oder bestimmte Interaktionsmuster, die mit der Konstruktion verbunden sind. Beispiel: die Sequenzmerkmale der teilspezifischen Konstruktion <verstehen mit Akkusativergänzung> sind «Bezug auf unmittelbaren Ko-Text; häufige Verwendung für Aktivitäten wie Bewerten und Argumentieren» (S. 283).
- Das Merkmal *Prosodie* wird nur dort aufgeführt, wo es die Konstruktion eindeutig bestimmt, jedoch nicht routinemässig bei allen Konstruktionen. Bsp. «fallende Intonation» (S. 41).

¹²<http://www.coli.uni-saarland.de/projects/salsa/> (30.11.2010)

Ein *weiteres Merkmal* <semType2> ist für die *Nomen* im Lexikon vorgesehen und soll den den *semantischen Typ* gemäss der CPA-Ontologie aufnehmen (z. B. [[Human]], [[Plant]], [[Abstract]] etc.). Das OLIF-Schema stellt zwar selber auch einen semantischen Typ für Nomina bereit (siehe Tabelle 9.4 auf Seite 216). Für die Auswertung der Verbpatterns (in der Analyseprozedur) ist es aber einfacher, den passenden semantischen Typ für die Nomen zusätzlich als Merkmal einzuführen als die beiden Ontologien zu matchen.

9.1.4 Weitere Überlegungen zum neuen Lexikon

Als Hinweis zu den Kategorisierungsbestrebungen mit den neuen OLIF-Feldern und zur Suche nach bedeutsamen Kategorien sei folgendes Zitat von Günthner und Bücken (2009) angeführt, das zeigt, dass nicht alle auftretenden Phänomene mühelos in Kategorien eingeordnet werden können:

Die Berücksichtigung auch peripherer Vertreter der jeweiligen Konstruktionen als vollwertige funktionstragende Einheiten in der Kommunikation zeigt ferner, dass das Postulieren diskreter Kategoriengrenzen der in der Interaktion beobachtbaren funktionalen und strukturellen Vielfalt grammatischer Konstruktionen nur unzureichend Rechnung trägt. Konstruktionen sind stattdessen eher im Sinne von (Schegloff et al., 1996, S. 40) als «contingent and concerted accomplishment, symbiotic with the setting of social interaction which is its home base zu betrachten». (Günthner & Bücken, 2009, S. 7)

Diese Bemerkung gilt natürlich nicht nur für den engeren Bereich der Konstruktion in der Grammatik, sondern auch für den weiteren Begriff der Konstruktion als umfassende Beschreibungsstruktur für sprachliche Phänomene. Günthner und Bücken (2009) plädieren sogar dafür, die Nicht-Kategorisierbarkeit als Herausforderung einer *realistischen Sprachwissenschaft*¹³ zu betrachten:

Als besonders fruchtbar kann sich hier ein Vorgehen wie das der «realistischen Sprachwissenschaft» erweisen, das analog zum konversationsanalytischen Prinzip <order at all points> (Sacks 1964–68/92) strukturelle und/oder semantischfunktionale Devianz nicht einfach als Randständigkeit oder Markiertheit deutet,

¹³Vgl. Auer (2003): «Realistische Sprachwissenschaft».

sondern ihr in Form der Frage «why that now?» (Sacks 1964–68/92) bzw. «welcher kommunikative Zweck liegt dem Abweichen vom Prototypen hier nachweislich zugrunde?» begegnet. (Günthner & Bückler, 2009, S. 15)

Diese Sichtweise führt ähnlich wie die Überlegungen zur *dynamischen Bedeutungskonstitution* über das Lexikon hinaus zu Kategorisierungsprozessen und Bedeutungszuweisungen, die sich individuell in der unmittelbaren Gesprächssituation entwickeln.

Nicht lexikalisierte Konstruktionen

Die Erweiterung des Lexikons läuft in einem zirkulären Entwicklungsprozess ab. Ausgangspunkt ist eine Konstruktion, die im Korpus gefunden wurde und die analysiert und ins Lexikon aufgenommen wurde. Diese Konstruktion wird im Lexikon mit den oben beschriebenen Formalismen abgebildet. Die Lexikoneinträge sollen Form-Bedeutungspaare mit möglichst spezifischen Eigenschaften sein, so dass es möglichst keine Konstruktionen gibt, die formal identisch beschrieben sind, die aber unterschiedliche Bedeutung haben. Die Darstellung der Konstruktion im Lexikon (d. h. des ganzen Lexikoneintrags) wird auf der andern Seite durch die Anforderungen der maschinellen Such- und Kodierprozeduren der Analyseapplikation bestimmt: Wie kann eine bestimmte Konstruktion im Text gefunden und von andern unterschieden werden?

Wo verläuft die Grenze zwischen *Grammatik* und *Lexikon*? Wo im System sollen Konstruktionen untergebracht werden, die keinen lexikalischen Anteil haben? Im vorliegenden Lexikonprojekt liegt der Fokus nicht auf diesen rein grammatischen Konstruktionen; trotzdem ist es nötig, die Integration dieser Konstruktionen zu festzulegen. Zuerst sollen noch einmal die Begriffe geklärt werden:

Lexikalisierte Konstruktionen: sind Ausdrücke, die als Ganzes im Lexikon aufgenommen sind wie z. B. die Konstruktion «den Vogel abschiessen». Obschon hier der Akteur variabel ist, soll diese Konstruktion als lexikalisiert bezeichnet werden. Auch eine Modifikation des Idioms («den grössten Vogel abschiessen»)¹⁴ soll hier noch möglich sein.

Teil-lexikalisierte Konstruktionen: Konstruktionen mit mindestens einem lexikalisierten Ausdruck und mindestens einem Ausdruck, der variabel ist und der keine prototypische Lexikalisierung aufweist (zum

¹⁴vgl. Keil (1997, S. 78).

Beispiel die Konstruktion $\langle \text{haben zu} + \text{INF} \rangle$). Diese Konstruktionen werden benannt auf Grundlage ihrer lexikalisierten Komponenten (die $\langle \text{haben-zu} \rangle$ -Konstruktion). Weiteres Beispiel für eine teillexikalisierte Konstruktion: $[\text{sich ADJx geben}]$, «den Eindruck erwecken, X zu sein», wobei X verschiedene Werte annehmen kann wie «stark», «cool», «freundlich» und weitere Adjektive¹⁵.

Nicht-lexikalisierte oder grammatische Konstruktionen: Konstruktionen, die aus linguistischen Strukturen bestehen und mit lexikalischen Items instanziiert werden können (Beispiel $\langle \text{Transitivkonstruktion} \rangle$, oder die Interrogativsatzkonstruktion $[\text{WH VP}]$). Das letzte Beispiel zeigt, dass die Abstraktion von Konstruktionen beliebig weit fortgesetzt werden kann.

Ziel des vorliegenden Lexikonprojektes ist es, möglichst viele Konstruktionen im Lexikon zu integrieren, die mindestens ein lexikalisiertes Element aufweisen. Für die Darstellung der *schematischen Konstruktionen* würde grundsätzlich die Möglichkeit bestehen, diese auch ins Lexikon aufzunehmen. Weil aber die schematischen Konstruktionen in der Regel aus syntaktischen Elementen bestehen und weil die syntaktische Beschreibung (mindestens für Konstruktionen auf der Satzebene) bereits im Satzmuster nach Wahrig im Lexikoneintrag vorhanden ist, sollen direkt diese Wahrig-Satzmuster als schematische Konstruktionsklassen verwendet werden.

Diese Überlegungen beziehen sich, wie gesagt, auf die Satzebene, die im vorliegenden Projekt stärkstes Gewicht erhält. Daneben gibt es schematische Konstruktionen auf andern Ebenen, die nicht mit den Satzmustern beschrieben werden können, wie zum Beispiel die bei Imo (2007) beschriebenen Konstruktionen $\langle \text{Modalwort} \rangle$, $\langle \text{Vergewisserungssignal} \rangle$ oder $\langle \text{Diskursmarker} \rangle$.

9.1.5 Darstellung der Konstruktionen

In diesem Kapitel wird beschrieben, wie die gefundenen Korpusbeispiele/Konstruktionen dargestellt und charakterisiert werden. Nach einer Beschreibung der Konstruktion, ihrer Verwendung und des Kontextes, und nach der Auflistung des Vorkommens in den verschiedenen Korpora wird die Konstruktion anhand der Eigenschaften aus dem JAKOB-Lexikon dokumentiert. Neben der Dokumentation der Konstruktion als solche ist der

¹⁵Diese Beispiele stammen aus dem Handout von Anatol Stefanowitsch, IDS-Tagung Mannheim, Februar 2010.

Hauptzweck dieser Darstellung eine Prüfung und Validierung der neuen Lexikonkategorien in Bezug auf die adäquate Vermittlung von Bedeutung mittels Konstruktionen. Die Darstellung im Lexikon kann mit den Merkmalslisten verglichen werden, wie sie z. B. Imo (2007, S. 40) für die Darstellung von Konstruktionen verwendet.¹⁶ Die zusammenfassende tabellarische Darstellung der Lexikoneinträge (siehe Tabelle 9.7) zeigt alle konstruktionsrelevanten Merkmale, die es ermöglichen sollen, verschiedene Konstruktionen zu disambiguieren.

Der *JAKOB-Code* wird in der tabellarischen Darstellung der Konstruktionen folgendermassen dargestellt: *Codekürzel* (der Code kann einfach oder aus zwei bis maximal drei Einzelcodes zusammengesetzt sein), darauf folgt die *Bezeichnung* des Codes, die dazu gehörige *Dimension* und die *Subdimension*, wie im Kapitel (2.3.4) vorgestellt. Die Definition der einzelnen Codes kann im Lexikon online durch Anklicken der Codelinks in den Lexikoneinträgen angezeigt werden.

Prosodische Merkmale werden in der aktuellen Version nicht verwendet. Dies hat damit zu tun, dass grosse Teile der Gesprächskorpora nur schriftlich (transkribiert) vorliegen.

9.2 Analysebeispiele, Konstruktionen

Die folgenden Konstruktionsbeispiele sind nicht systematisch geordnet. Ein Versuch der Einordnung der Konstruktionen in Gruppen könnte z. B. nach den folgenden Kriterien vorgenommen werden:

Kommunikationskonstruktionen bilden besondere Arten der Kommunikation, aber auch der Situation und des Verhaltens in der Situation ab. Diese Konstruktionen beinhalten auch Aspekte der *Positionierung*, die ebenfalls als ein Gruppenkriterium betrachtet werden könnte. Eine grosse Gruppe von Konstruktionen dient dazu, *Emotionen* auszudrücken, z. B. *Angst haben*. Der Ausdruck von Emotionen geschieht hauptsächlich figurativ, z. B. durch die Verwendung von Metaphern wie *explodieren* oder *Dampf ablassen* (Behältermetapher). *Somatismen* nehmen einen wichtigen Platz in der Konstruktionstypologie ein. Eine weitere Gruppe von Konstruktionen beinhaltet Ausdrücke mit Tieren und Tierbezügen.

Daneben gibt es auch formale Einteilungskriterien, wie voll lexikalisierte vs. teillexikalisierte Konstruktionen (z. B. <typisch Frau/Mann>). Dies als Beispiel für eine teillexikalisierte Konstruktion, die sich wegen ihrer Einfachheit gut für die illustrative Darstellung von Konstruktionen eignet. Funkti-

¹⁶Darstellung der Konstruktionen in Wert-Attribut-Kästen (Imo, 2007, S. 40).

Tabelle 9.7: Konstruktion: Musterdarstellung

Attribut	Wert	Bemerkung
canForm	kanonische Form	Lemma, Grundform, Nennform
orthVariant	orthographische Variante	(ohne Bedeutungsunterschied)
crossRef	Cross Reference, Verknüpfung	verschiedene Linktypen
ptOfSpeech	Wortart	(Wortart des Heads)
head	Kopf des MWE	gemäss Konvention am Anfang des Ausdrucks
entryFormation	Art des Eintrags	u.a. single word vs. multi word
phraseType	Typ des MWE	falls entryFormation = phrase
synFrame	Satzmuster für Verben	syntaktisches Satzmuster nach Wahrig
synType	syntaktischer Typ	z. B. Haupt-, Modal, Funktionsverb
semType	semantische Typen	gemäss OLIF-Ontologie
definition	freie Beschreibung	Paraphrase, Bedeutungserklärung
natclause	Mustersatz	Eingabe für Parser
bauplan	Struktur für NLP	maschinenlesbare Darstellung der Konstruktion
pattern	Verbpattern	semantisches Satzmuster (für Verben)
construction	Konstruktion (syntaktisch)	gemäss CxG
subjField	Domain, Genre	Diskurstyp
textType	linguistische Marker	Diskurseinheit
register	lokale Stilmerkmale	lokale Diskursmuster
topic	lokales Diskursthema	lokales Pattern
Funktion	Kommunikative Funktion	noch nicht im Lexikon enthalten
JAKOB-Code	Kodierung	Code – Bezeichnung – Dimension – Subdimension

onsverbgefüge können ebenfalls unter einem formalen Aspekt als Konstruktionen betrachtet werden.

9.2.1 Schematische Konstruktion: <lesen>

Zuerst soll eine schematische Konstruktion vorgestellt werden, nämlich die *Transitivkonstruktion* am Beispiel des Verbs «lesen», wie sie in den drei Korpora der Psychotherapiegespräche (WIL, AMA, GUS) vorkommt. Zuerst wird das Vorkommen des Verbs «lesen» gezählt (Tabelle 9.8).

Lesen steht hier stellvertretend für die Klasse der Transitivekonstruktion, die wenig Spezifisches über das Verb aussagt. Das Verbpattern geht

Tabelle 9.8: Konstruktion <lesen>

Konstruktion	WIL	AMA	GUS	AGD	CHTK	deWAC
lesen (Gesamtzahl)	313	306	25	1062	4469	374'565

aber über die Transitivitykonstruktion hinaus und spezifiziert die Argumente näher: *Lesen* kann transitiv verwendet werden, kann aber auch mit einem zusätzlichen Argument verwendet werden. Beide Argumente sind fakultativ, denn *lesen* kann auch ohne Argument verwendet werden («ich lese»). Die Bedeutung des Verbes *lesen* ist in diesen drei Varianten sehr ähnlich, so dass die Varianten gemeinsam in einem Lexikoneintrag abgebildet werden können. Tabelle (9.9) zeigt den Lexikoneintrag für <lesen>.

Tabelle 9.9: Konstruktion: <lesen>

Attribut	Wert	Bemerkung
canForm	<i>lesen</i>	
crossRef	überfliegen, verschlingen	near-synonym, Metapher
ptOfSpeech	Verb	
head	lesen	
entryFormation	sgl	Einzelwort
synFrame	505	Verb (+ AkkO) (+ PräpO)
synType	main-vb, transitiv	
semType	act, percept	unspezifische Aktivität, FrameNet: sentient
definition	lesen, ein Buch lesen (über etwas), über etwas lesen	
FrameNet	<i>Reading</i> : The Reader attends to a Text to process its information.	
Dornseiff	12.54 Lesen (Wahrnehmung, Verstehen, Lernen)	
natclause	jmd liest etwasn über etwasn	
bauplan	500: 1 - - L2 subj Human 2 lesen VVFIN L0 root n/a 3 - - L2 obja Document 4 über APPR L2 pp n/a 5 - - L4 pn Anything	
pattern	[[Human]] lesen ([[Document]]) ({über [[Anything]])}	
construction	Transitivkonstruktion [NP _{nom} V _{fin} NP _{akk}]	
subjField	Therapiegespräch	
textType	unspezifisch	Diskurseinheit
register	unspezifisch	lokale Diskursmuster
topic	information	
Funktion	Informationsbeschaffung, Transitivityschema	
JAKOB-Code	TUN	TUN=machen (Geschehen – Vorgehen)

- Die drei erwähnten Varianten werden im selben Lexikoneintrag repräsentiert: *lesen* ohne Objekt, *ein Buch lesen*, *ein Buch über etwas lesen*, *ein Buch zu einem Thema lesen*.
- Weitere freie Ergänzungen zur Konstruktion/zum Verbpattern sind möglich: *ein Buch am Abend lesen*.
- Eine vierte Variante der Konstruktion ist *lesen mit einem Nebensatz*: («ich habe dann noch gelesen, dass der Hypothalamus oder wie das Ding heisst...», Korpus WIL). Strukturen, die über den Satz hinausgehen, werden zum gegenwärtigen Zeitpunkt nicht im Lexikon abgebildet.
- Mit dem Verb *lesen* gibt es ausserdem viele Phraseologismen: *Gedanken lesen*, *von den Lippen/Augen (ab)lesen*, *im Chor lesen* («die Kinder lesen im Chor, der Lehrer schreibt ein Wort an die Tafel», CHTK), *im Chor sprechen* (AGD). Diese Phraseologismen müssen dann ins Lexikon aufgenommen werden, wenn ihre Bedeutung nicht aus den Einzelkomponenten, wie sie im Verbpattern aufgeführt sind, erschlossen werden kann. *Gedanken lesen* müsste dann als verfestigter Phraseologismus (set-phrase) betrachtet werden, ebenso *im Chor lesen*.

Aus den Häufigkeitswerten in den Korpora geht möglicherweise hervor, dass Amalie wesentlich häufiger liest als Wilma; Amalie braucht das Verb *lesen* 306 mal (bei einer Gesamttokenzahl von 1.1 Mio), Wilma verwendet das Verb 313 mal (Gesamttokenzahl: 2.4 Mio).

9.2.2 Kommunikationskonstruktion: <an den Karren fahren>

Ausgangspunkt für die Untersuchung dieser Gruppe von Konstruktionen mit dem Bestandteil «Karren» ist die Konstruktion <an den Karren fahren> aus dem Korpus WIL.

- (54) und da sind Sie ja auch der Kollegin ein bisschen an den Karren gefahren... (WIL.259.330)

Der Therapeut paraphrasiert ein Verhalten der Klientin, die sich im Geschäft gegenüber einer Kollegin zur Wehr setzt, mit dem umgangssprachlichen metaphorischen Ausdruck «an den Karren fahren». Diese Konstruktion

kommt bei Wilma 13 mal vor und wird mit einer Ausnahme nur vom Therapeuten verwendet, der damit eine Art Formel für das «Sich-Behaupten» etabliert. In den beiden Korpora AMA und GUS kommt diese Konstruktion nicht vor; bei Amalie wird «Karren» viermal als Bezeichnung für ein Auto verwendet und einmal in der Konstruktion <sich vor den Karren spannen lassen>.

Gesprächsausschnitte Korpus WIL

(55) Wilma: an den Karren fahren 1 (WIL.259.330)

- 01 T: hmhm hmhm +
 02 P: solche Sachen machte
 03 weil + regieren durfte ich auch nicht
 04 T: hör auf zu regieren + und regieren
 05 die Regierung war ja die Mutter
 06 P: ja
 07 T: das heisst doch
 08 hör auf so zu sein wie die Mutter
 09 P: nein ich denke es hiess etwas anderes
 10 T: was denn?
 11 P: hör auf der Mutter ins Gehege zu geraten
 12 T: aha ja also die quasi zu konkurrenzieren
 13 P: hmhm
 14 T: hmhm hmhm und da sind Sie ja auch der Kollegin ein bisschen
 an den Karren gefahren +
 und Sie haben das auch
 16 P: habe ich heute ganz offensichtlich gemacht +
 17 T: und Sie signalisierten das auch äh durch so verschiedene
 Zeichen auch + signalisiert
 18 P: ja +
 19 T: aber das hat sie gar nicht so irritiert
 20 es war auch gar nicht negativ und wurde auch von den anderen
 gar nicht negativ aufgefasst im Gegenteil,
 21 offensichtlich ganz wichtige Fragen
 22 die geklärt werden mussten
 23 P: es irritierte mich zwischendurch schon.
 24 ich liess es einfach ganz bewusst ausser acht
 25 T: hmhm ja ja
 26 P: oder ?
 27 T: ja

Wilma erzählt von einer geschäftlichen Sitzung, bei der die übrigen Sitzungsteilnehmer schlecht vorbereitet waren; sie hatte als einzige die Unterlagen bei sich, und der Sitzungsleiter musste sich bei ihr orientieren. Dadurch hatte Wilma das Gefühl, die Sitzung zu *dominieren*. Auf die Nachfrage des Therapeuten, warum sie dies als negativ und als Befürchtung formuliere,

sagt Wilma, dass sie bereits in der Kindheit oft in die Schranken verwiesen wurde und dass sie nicht «regieren» durfte, weil «die Regierung» die Mutter innehatte. Der Therapeut versteht nun, dass Wilma sich in der erzählten Situation selbst verbieten möchte, so wie die Mutter zu sein, die Patientin nimmt jedoch diese Interpretation nicht an, sondern sagt, dass es darum gegangen sei, die Mutter nicht zu konkurrenzieren, ihr nicht «ins Gehege zu geraten». Dieser von der Klientin vorgeschlagene idiomatische Ausdruck wird jedoch vom Therapeuten nicht aufgenommen, sondern er formuliert eine aktive Variante «jemandem an den Karren fahren». Diese Ausdrucksweise wird aber wiederum von Wilma nicht aufgenommen; auch in den weiteren Stunden wird der Ausdruck – mit einer Ausnahme – nur vom Therapeuten gebraucht.

Wie wird die Konstruktion verwendet? Der Therapeut umschreibt Wilmas Aktion gegenüber der Kollegin als *an den Karren fahren* und braucht damit einen idiomatischen Ausdruck, der Interpretationsspielraum lässt und durch eine gewisse umgangssprachliche Rauheit ausdrückt, dass Wilma richtig gehandelt hat, indem sie sich verteidigte und nicht unterkriegen liess. Diese Sicht wird denn auch von Wilma mit einem gewissen Stolz bestätigt («habe ich heute ganz offensichtlich gemacht», Z.15).

Weitere Beispiele «an den Karren fahren»

Im Folgenden werden weitere Beispiele (hier nur für das Korpus WIL) vorgestellt und kurz kommentiert. Die Konstruktion wird 12 mal vom Therapeuten gebraucht, von Wilma nur einmal.

Wilma erzählt von der Mutter, für die sie nur Statistin ist für deren Selbstdarstellung; das gleiche in Bezug auf die Schwester. Der Therapeut verwendet *an den Karren fahren* mehrmals, um die Angst der Klientin zu umschreiben, sich gegenüber Mutter und Schwester zu behaupten und mit einem gewissen Nachdruck und Aggressivität ihre Position zu verteidigen (Beispiel 56).

(56) Wilma: an den Karren fahren 3 (WIL.156.523)

```

01 P: und bei §Regine ich glaube
02   was was eben auch noch ein Stück + weit eine Rolle spielt
03   ist die Tatsache
04   dass §Regine eigentlich äh , noch noch (03) der einzige Faden ist ,
05   den ich wirklich als Faden wahrnehmen kann innerhalb dieser Familie
    jetzt zu mir
06 T: hmhm
07 P: oder ?

```

- 08 T: und es wäre dann wie , der letzte Faden ,
 noch zu gefährden oder vielleicht sogar zu verschneiden
 09 wenn Sie ihr so an den Karren fahren würden

Der Therapeut regt wiederholt an, dass Wilma sich wehren und aktiver ihren Standpunkt verteidigen sollte, und ermutigt sie, z. B. der Mutter an den Karren zu fahren. Auch in der therapeutischen Situation wird die Analogie aufgegriffen, und der Therapeut fragt am Ende der Stunde 199:

(57) Wilma: an den Karren fahren 4 (WIL.199.348)

- 01 T: könnte das nicht etwas damit zu tun haben
 02 dass Sie dann nachher das Gefühl hatten
 03 die Mutter will nicht dass ich ihr ebenbürtig oder überlegen wird
 04 und dass das vielleicht jetzt auch ein bisschen bei uns
 ein wenig drohte
 05 dass es vielleicht auch für sie der Grund war
 06 so sehr zurückhaltend zu sein mit der Psychologie
 07 so dass ich erst gar nicht erst verstand
 08 worum es da dass es da um Psychologie geht auch !
 09 also dass da vielleicht noch ein bisschen mitschwingt ja :
 10 dass Sie das Gefühl haben
 11 ja Sie würden mir irgendwie an den Karren fahren
 12 P: (lacht)
 13 T: und ich ich ich ich möchte nicht
 14 dass sie das Gleiche machen oder etwas Ähnliches machen
 15 und möchte Sie quasi wie die Mutter , schön unten halten
 16 also ich bin der
 17 der die Kontrolle hat und die Übersicht und so
 18 und sonst aber gar niemand.

Später (WIL.272.186) erzählt Wilma von einem Nachtesen mit einer Kollegin und dass ihr diese vorgeworfen habe, Wilma sei neuerdings viel bestimmter und aggressiver, und ob sie, die Kollegin, wohl als nächste ebenfalls «drankomme» (aggressiver behandelt werde). Das bestimmtere Auftreten und die «Aggressivität» werden als Fortschritte in der Therapie gewertet, und der Therapeut findet es richtig, dass Wilma ihre Meinung nicht zurückhält und der Kollegin auch mal *an den Karren fährt*, wenn dies notwendig ist.

Komplementär zur Konstruktion *an den Karren fahren* verwendet der Therapeut die Konstruktionen *ins Gärtchen trampen* und *einen riesigen Haufen schlucken*, um die Zusammenhänge zu paraphrasieren: Aus Angst davor, der Mutter ins Gärtchen zu trampen, fährt sie ihr nicht an den Karren und muss dafür einen riesigen Haufen schlucken:

T: ...ich denke dass Sie die Schwächen der Mutter sehr genau wahrnahmen und wahrscheinlich merkten dass die durcheinander kommen würde Angst haben würde wenn Sie ihr *an den Karren fahren* würden dass Sie da gar nicht so souverän ist wie sie tut und ich glaube Sie hielten sich aus dem Grund zurück aus dem Grund *sind Sie ihr nicht ins Gärtchen getrampt*, mussten da natürlich *einen riesigen Haufen schlucken* aus dem Grund.
(WIL.272.196)

Ins Gärtchen trampen,¹⁷ wie auch die weiter oben erwähnte Konstruktion *ins Gehege geraten* können als komplementäre Varianten von *an den Karren fahren* betrachtet werden, die mit weniger aggressiver Absicht assoziiert sind als die aktive Variante *an den Karren fahren*. Die Konstruktion *einen riesigen Haufen schlucken* kann als Beispiel einer nicht lexikalisierten Metapher angesehen werden (die Kombination kommt in keinem der verwendeten Korpora vor, ausser an der oben zitierten Stelle).

(58) Wilma: an den Karren fahren 5 (WIL.281.62)

01 P: also eben als ich sie dann fragte äh
02 ob sie sich jetzt da deswegen angegriffen gefühlt hat
03 das hätte ich doch früher nicht gemacht
04 T: hm
05 P: um Gottes Willen auch !
06 T: hmhm
07 P: ich wäre gar nicht auf die Idee gekommen
08 T: dass es so ist +
09 P: sondern dass + ich dem anderen äh jetzt wirklich an den Karren
gefahren wäre und hätte versucht mich zu entschuldigen
10 und ich weiss nicht was bei Situation wo das gar nicht nötig ist ,
11 und äh , ja , mich stört das nicht
12 wenn wir ab und zu lauter miteinander diskutieren also ,
13 man kann mit mir eben auch so reden und ,
14 im Unterschied zu früher ist es eben so
15 dass ich äh früher zugelassen habe
16 dass man mit mir so redet ,
17 aber ich hätte nie Gleiches mit Gleichen vergolten hätte ja
18 und heute mache ich das einfach ziemlich diskussionslos
19 aber ich musste mich bis jetzt nicht einmal richtig entschuldigen

Der Therapeut ermutigt die Klientin in mehreren Anläufen, sie dürfe ihren Gesprächspartnern *an den Karren fahren*. Wilma beherzigt dies auch

¹⁷ «trampen»: gemäss Schweizer Wörterbuch (Meyer & Bickel, 2006): schwerfällig treten (mundartnah). Auch im CHTK sind 4 Fundstellen dokumentiert. Nicht zu verwechseln mit dem Verb «trampeln» (929 Treffer im deWAC).

und erzählt in mehreren Episoden von ihren Erfolgen mit dieser «neuen Strategie». In der Episode (WIL.281.62, Beispiel 58) braucht Wilma selber die Konstruktion *an den Karren fahren*, als sie von einer Auseinandersetzung mit Geschäftskollegen erzählt. In der letzten Episode (WIL.289.488) wird noch einmal die Therapiesituation thematisiert, indem der Therapeut betont, dass das *an den Karren fahren* auch gegenüber ihm in der Interaktion wichtig sein könnte:

T: ich denk- ich denke auch noch ein bisschen an der an diese Vorsicht von ihnen . *Fahren sie mir nicht zu fest an den Karren* was ist da mit meiner Expertenschaft äh: äh ja dass ich dann irgendwie gekränkt oder beleidigt oder eingeschnappt, es könnte natürlich so ein bisschen dazu führen mit solchen Dingen vorsichtiger zu sein, obwohl es vielleicht etwas ganz Wichtiges wäre. (WIL.289.488)

Zusammenfassend kann festgehalten werden, dass die Konstruktion <an den Karren fahren> in den Transkripten WIL 13 mal verwendet wird, hauptsächlich vom Therapeuten, und dass damit immer die gleiche Bedeutung gemeint wird. Die metaphorische Redensart ermöglicht Vagheit in der Bedeutungskonstitution und kann so als Formel eingesetzt werden, die in verschiedenen Situationen angewendet werden kann. Die aktiven Konstruktionen *an den Karren fahren* und *ins Gärtchen trampen* werden fast ausschliesslich vom Therapeuten gebraucht und von der Klientin nicht übernommen; sie verwendet dagegen die passive Konstruktion *ins Gehege geraten*. Tabelle (9.10) zeigt die Konstruktion, wie sie im Lexikon festgehalten wird.

Konstruktion: <den Karren ziehen>

Eine weitere Konstruktion mit dem Bestandteil *Karren* kommt im Korpus WIL einmal vor (vom Therapeuten gesagt): *den Karren ziehen*.

(59) Wilma: den Karren ziehen (WIL.253.39)

- 01 P: alles da (lachen leicht) ja ja ; ja und jetzt weiss ich nicht
mehr weiter
02 T: hola , jetzt stehen wir da
03 P: ja
04 T: wie der Esel am Berg (lachen)
05 P: ja warum auch nicht
06 wobei äh ich finde sonst Esel nicht so schön (lachen)
07 ja richtig
08 mir kam gerade so , in den Sinn

Tabelle 9.10: Konstruktion: *an den Karren fahren*

Attribut	Wert	Bemerkung
canForm	<i>fahren an den Karren</i>	
crossRef		
ptOfSpeech	Verb	
head	fahren	
entryFormation	phrase	
phraseType	Idiom, Metapher	
synFrame	Satzmuster 605	Dativobjekt und Präpositionalobjekt
synType	main-vb	
semType	act	unspezifische Aktivität
definition	jmdm. durch seine Äusserungen, Handlungen schaden wollen (Duden 11), jmdm zu nahe treten, jmdn grob beleidigen (Wahrig)	
natclause	jmd fährt jmdm an den Karren	
bauplan	605: 1 - - L2 subj Human 2 fahren VVFIN L0 root n/a 3 - - L2 objd Human 4 an APPR L2 pp n/a 5 die ART L6 det n/a 6 Karren NN L4 pn n/a	
pattern	[[Human]] fahren [[Human]] {an den Karren}	
construction	[an den Karren fahren]	
subjField	Therapiegespräch	Diskurstyp nicht festgelegt
textType	unspezifisch	Diskurseinheit
register	umgangssprachlich, salopp	lokale Diskursmuster
topic	Sich Durchsetzen	am Arbeitsplatz und gegenüber der Mutter
Funktion	Kommunikation mit aggressiver Selbstbehauptung, Positionierung	
Modifikationen	(Duden 11) <i>an den Karren pinkeln</i>	
JAKOB-Code	SIG-ATT	SIG=Sprechen (Geschehen – Vorgehen), ATT=Angreifen (Geschehen – Vorgehen)

09 dass ich nicht mehr weiter weiss

10 T: ja wie wenn Sie da jetzt den Karren ziehen müssten

«Ja, wie wenn Sie da jetzt den Karren ziehen müssten»:¹⁸ Der Therapeut verwendet diesen Ausdruck in der Stunde 253 und gibt der Klientin damit

¹⁸Original: «de Charre schläike». Hier scheint eine minimale Übersetzung angebracht, das das Verb «schleiken» oder «schläiken» in keinem Wörterbuch oder Korpus vorkommt. Im Internet (Google-Suche) finden sich wenige Belege in der ch-Domain: schleiken ist ein Ausdruck aus der Fischerei (Fischen mit Schleppnetzen), und auch Holz kann geschleikt werden.

eine *candidate answer* oder *assessment/sentence completion* (Antaki, 2007, S. 533), als sie feststellt, dass sie im Gespräch nicht mehr weiter weiss. Die Klientin hat vorher (zu Beginn der Sitzung) erzählt, wie sie im Wartezimmer des Therapeuten im Laufe der Zeit eine ebenfalls wartende Klientin kennen gelernt hat, wie sie mit ihr ins Gespräch gekommen ist und wie sie schliesslich sogar einen gemeinsamen Museumsbesuch mit anschliessendem Spaziergang gemacht haben. Wilma findet es «interessant», was sich «aus dem Wartezimmer entwickeln kann». Nach diesem Bericht weiss sie nicht mehr weiter, und der Therapeut sagt: «Jetzt stehen wir da wie der Esel am Berg». Gemäss Antaki (2007) stellt die Verwendung von Idiomen durch den Therapeuten auf den ersten Blick Nähe her. Durch die informelle Ausdrucksweise gibt sich der höher gestellte Therapeut quasi zum Klienten «hinunter». Durch die Verwendung von clichéhaften Idiomen wird aber gleichzeitig das Gespräch anders evaluiert. Der idiomatische Ausdruck ist «universal» und für jeden Zweck verwendbar; der Therapeut kann so eine Art abschliessendes Statement machen, das nicht mehr so einfach hinterfragt werden kann (Antaki, 2007, S. 529), siehe Tabelle (9.11).

Korpusvorkommen: Konstruktionen mit *Karren*

In der Tabelle (9.12) sind die Häufigkeiten der verschiedenen Konstruktionen mit «Karren» aufgeführt. Es folgen einige Bemerkungen zu den Korpuszahlen:

- Im Korpus *deWAC* kommt der Begriff *Karren* 3598 mal vor. Beste *Kollokationskandidaten* zu *Karren* sind (in Rangreihenfolge gemäss *logDice-Score*):¹⁹ *Dreck* (257/9.1), vgl. «*Karren* aus dem *Dreck* ziehen» oder «der *Karren* steckt im *Dreck*»), *spannen* (459/9.1, «sich vor den *Karren* spannen lassen», «vor den *Karren* spannen»). Das Verb *fahren* erscheint als *Kollokationskandidat* erst ca. auf dem 50. Rang, trotz hoher Frequenz (252/4.5). Die Konstruktion *an den Karren fahren* kommt in dieser Auswahl 62 mal vor.
- Im *CHTK* kommen die Konstruktionen *den Karren ziehen* und *den Karren aus dem Dreck ziehen* vor. Weitere Konstruktionen aus dem *CHTK* sind: *der Karren läuft*, *den festgefahrenen Karren wieder in Schwung bringen*, *der Karren ist verfahren* (festgefahren), *der Karren steht am Berg*, man sollte *den Karren nicht überladen*.

¹⁹Die beiden Zahlen bedeuten: *Häufigkeit* der Kookkurrenz (des gemeinsamen Auftretens) und *logDice-Score* (siehe Kapitel 5.1.2).

Tabelle 9.11: Konstruktion: *den Karren ziehen*

Attribut	Wert	Bemerkung
canForm	<i>ziehen den Karren</i>	
crossRef	schleppen den Karren (synonym)	übernehmen die Initiative (near-synonym)
ptOfSpeech	Verb	
head	ziehen	
entryFormation	phrase	Phraseologismus
phraseType	Idiom, Metapher	
synFrame	500	Verb + AkkO
synType	main-vb	
semType	act	unspezifische Aktivität
definition	die Initiative übernehmen	
natclause	jmd zieht den Karren	
bauplan	500: 1 - - L2 subj Human 2 ziehen VVFIN L0 root n/a 3 die ART L4 det n/a 4 Karren NN L2 obja n/a	
pattern	[[Human]] ziehen {den Karren}	
construction	[den Karren ziehen]	
subjField	Therapiegespräch	
textType	unspezifisch	(Diskurseinheit)
register	umgangssprachlich	
topic	Verantwortung übernehmen	
Funktion	Bechreibung initiativen Handelns	
JAKOB-Code	ARB-WOR	ARB=engagieren (Handeln – Interagieren), WOR=Verantworten (Wollen – Festlegen)

Tabelle 9.12: Konstruktionen mit *Karren*

Konstruktion	WIL	AMA	GUS	AGD	CHTK	deWAC
an den Karren fahren	13	0	0	0	0	62
den Karren ziehen	1	0	0	0	3	33
den Karren aus dem Dreck ziehen	0	0	0	0	1	75
den Karren schleppen	1	0	0	0	0	4
sich vor den Karren spannen lassen	0	1	0	0	0	129
vor den Karren spannen	0	0	0	0	0	98
den Karren in den Dreck fahren	0	0	0	0	0	44
Karren (Nomen)	15	5	0	52	125	3598

- Im *AGD* und im *CHTK* existieren keine Belege für *an den Karren fahren*.
- Im Korpus *deWAC* kommt der Ausdruck «an den Karren» 92 mal vor. Die besten *Kollokationskandidaten* zu diesem Ausdruck sind: an erster Stelle *pinkeln* (4/5.8), an zweiter Stelle *fahren* (65/2.5), an dritter Stelle folgt *spannen* (4/2.49).²⁰
- Im *Korpus AMA* kommt nur eine Konstruktion mit *Karren* vor, nämlich sich vor den Karren spannen lassen: «...das kann man tatsächlich, nämlich wenn die anderen sich vor den Karren spannen lassen»
- *karren* kommt auch als Verb vor (887 Treffer im *deWAC*).

Verwandte Konstruktionen

Die Konstruktion *jemandem ins Gehege geraten/kommen* wurde oben schon erwähnt und bildet eine Art Komplement zur Konstruktion *an den Karren fahren*. Wer jemandem an den Karren fährt, gerät ihm ins Gehege. Diese Konstruktionen kommen im Korpus *Wilma* achtmal vor, «Der Mutter ins Gehege geraten» sagt der Therapeut dreimal, «ins Gehege kommen» wird von *Wilma* 4 mal verwendet, vom Therapeuten einmal. Im Korpus *deWAC* erscheint *ins Gehege geraten/kommen* 265 mal (9/256), im *CHTK* einmal, im *AGD* dreimal.

Die aktive Konstruktion *in den Garten/ins Gärtchen trampen* wird vom Therapeuten 10 mal verwendet. Die Konstruktion *den Laden schmeissen* hat eine ähnliche Funktion wie *den Karren ziehen*:

(60) *Wilma*: den Laden schmeissen (WIL.068.616)

```

01 T: wie wenn Sie alles alleine machen müssten
02 P: hmhm
03 T: gewissermassen
04 P: hmhm
05 T: den Laden alleine schmeissen
06 P: hmhm
07 T: das , dünkt mich ja nicht ! verwunderlich ,
08     oder was Sie nicht alles von Verantwortung schon alles ,

```

²⁰Mit diesen drei Werten wird sehr informativ die Funktion *Kollokationskandidaten* aus der *sketch engine* demonstriert: die besten Kollokationskandidaten sind nicht die häufigsten, sondern die mit dem salientesten gemeinsamen Vorkommen.

9.2.3 Konstruktion ‹Sorge tragen›

(61) Wilma: Sorge tragen 1 (WIL.085.222)

- 01 P: es prägte sicher das Bild ja
 02 T: und also , was ist denn das jetzt für ein Bild ?
 03 das ist ja ein Mann
 04 der sich hinter dem Rockzipfel verbirgt , oder ? der Frau ?
 05 P: hmhm
 06 T: im Grunde genommen ein diffiziler
 wie Sie sagen eher , schwacher Mann (p:04)
 07 verletzbar auch irgend + wie
 08 P: ja +
 09 T: oder ?
 10 P: hmhm
 11 T: zu dem man , Sorge tragen muss (wo mä muäss Sorg gä) (p:06)
 12 P: ja man kann dem so sagen (p:05)
 13 äh ich würde es vielleicht noch anders formulieren
 14 auf den man immer Rücksicht nehmen + muss
 15 T: ja + ja
 16 P: + so
 17 T: genau + (p:04)

Im Vorfeld der zitierten Episode spricht Wilma über ihren Vater und ihr Männerbild, dass durch den Vater mitbestimmt wurde, der einerseits aufbrausend war und schnell einen Wutanfall bekam, aber auf der andern Seite auch etwas Zerbrechliches an sich hatte, was durch seine Hörbehinderung akzentuiert wurde. Wilma bezeichnet es als «Rücksicht nehmen», was der Therapeut mit «Sorge tragen» bezeichnet.²¹

Das erste Beispiel zeigt die Konstruktion *Sorge tragen* in der Funktion «sich kümmern um einen Menschen», «Rücksicht nehmen auf einen Menschen». Diese Bedeutung entspricht nicht der des Verbs *sorgen*.

(62) Wilma: Sorge tragen 2 (WIL.173.046)

- 01 P: und er meinte dann
 02 äh im Moment lasse nichts auf ein Schleudertrauma schliessen
 03 aber äh , ich müsse es halt einfach beobachten
 04 man müsse halt allenfalls ein wenig Sorge tragen , oder ?
 05 und ich bin ja beim Auto fahren kannst du dich so drehen , oder ?

Der zweite Gesprächsausschnitt stammt aus einer längeren Schilderung Wilmas über die Folgen eines Autounfalls, mit einer Serie von Arztbesuchen,

²¹ *Sorge tragen* ist eine Übersetzung, die aber im wesentlichen die Bedeutung des Mundartausdrucks *Sorg gä* (Sorge geben) trifft. Die gebräuchlichen syntaktischen Möglichkeiten des Mundartausdrucks sind «ich gebe Sorge zu einem Gegenstand, auch zu einer Person», und «ich gebe ihm/mir Sorge».

auftretenden Komplikationen und verschiedenen Therapien, eine so genannte *Karriereerzählung* im Sinne der Erzählanalyse JAKOB. Ein Befund in der Folge des Unfalls ist ein mögliches Schleudertrauma, das jedoch nicht eindeutig diagnostiziert werden kann. Sie bekommt vom Arzt die Empfehlung, die Beschwerden zu beobachten und sich Sorge zu tragen, in der Formulierung von Wilma «man muss ein wenig Sorge tragen», zweimal mit der Modalpartikel *halt* kombiniert.

Die Funktion des Ausdrucks *Sorge tragen* ist hier vergleichbar mit der oben erwähnten «Rücksicht nehmen auf einen Menschen». Da es in diesem Fall um die eigene Person geht, wird die Rücksichtnahme abgeschwächt («ein wenig», «halt», «einfach»), wie es die gesellschaftliche Konvention erfordert.

(63) Wilma: *Sorge tragen* 3 (WIL.165.379)

- 01 T: also es geht ja da im Grunde genommen ganz fest darum
 02 dass Sie sich jetzt auch können und es auch machen sich schützen,
 + oder ?
 03 P: hmhm +
 04 T: Sie sagen es ja , sehr klar
 05 Sie sagen
 06 ich gehe nicht mehr die Ohrfeigen holen
 07 P: hmhm
 08 T: also sich wirklich zu sich schauen und sich *Sorge tragen*
 09 P: ja
 10 T: und sich schützen auch + also
 11 P: hmhm +
 12 T: und dann klar kontern
 13 und dann irgendwie in eine Angriff- eine Angriffsfläche bieten,
 oder ?
 14 wo die Mutter gerade wieder äh , loslegen kann

Im dritten Gesprächsausschnitt geht es darum, dass sich Wilma von der Mutter abgrenzen kann. Nachdem sie immer wieder in Situationen gerät, in der die Mutter in bevormundender Weise in ihre Entscheidungen eingreift, sagt der Therapeut, es sei wichtig, zu sich zu schauen und *sich Sorge zu tragen*. Hier wird die Bedeutung der Konstruktion *Sorge tragen* ergänzt mit Aspekten der Selbstachtung und Selbstverwirklichung. Tabelle (9.13) zeigt die Konstruktion <haben Sorge zu jmdm/etwas>:

Im ersten Beispiel (61) ist die Konstruktion *Sorge tragen* mit der Präposition *zu* realisiert: «Sorge tragen zu jemandem», im zweiten Beispiel (62) ohne Präpositionalobjekt: «Sorge tragen». Im dritten Gesprächsausschnitt (63) wird die Konstruktion reflexiv verwendet: «sich Sorge tragen». Es gibt eine weitere häufige Konstruktion, die in den drei Therapiekorpora nicht erscheint: «Sorge tragen für etwas».

Tabelle 9.13: Konstruktion *Sorge tragen zu jmdm/etwas*

Attribut	Wert	Bemerkung
canForm	<i>tragen Sorge [zu etwas]</i>	
crossRef		
ptOfSpeech	Verb	(head)
head	tragen	
entryFormation	phrase	
phraseType	set-phrase	Phraseologismus
synFrame	505	Verb + AkkO + (PräpO)
synType	funct	Funktionsverb
semType	act	unspecified activity
definition	sorgfältig umgehen, sich kümmern um	
natclause	jmd trägt Sorge zu etwasn	
bauplan	505: 1 - L2 subj Human 2 tragen VVFIN L0 root n/a 3 Sorge NN L2 obja n/a 4 zu APPR L2 objp n/a 5 - L3 pn Anything	
pattern	[[Human]] tragen Sorge (zu [[Anything]])	
construction	[tragen Sorge]	
subjField	Therapiesgespräch	Diskurstyp
textType	unspezifisch	Diskurseinheit
register	unspezifisch	lokale Diskursmuster
topic	Fürsorge, Pflege	Thema
Funktion	Rücksichtnahme, sorgfältiger	Umgang, Beschützen
JAKOB-Code	SOR	SOR=bangen (Wollen – Intendieren)

Korpusvorkommen: Konstruktionen mit *Sorge*

Die Tabelle (9.14) zeigt die Häufigkeiten der Konstruktionen mit «Sorge» in den Korpora.

Tabelle 9.14: Konstruktion «Sorge tragen»

Konstruktion	WIL	AMA	GUS	AGD	CHTK	deWAC
Sorge tragen zu	3	0	0	0	15	
Sorge tragen für	0	0	0	3	22	1161
Sorge tragen + DatO	0	0	0	0	13	
Sorge tragen	0	0	0	0	10	
Sorge (Nomen)	28	14	2	111	1396	50652

Im Schweizer Wörterbuch (Meyer & Bickel, 2006) ist die Konstruktion so aufgeführt: «[zu] jmdm/etwas Sorge tragen», ausserdem wird eine zweite *mundartnahe* Variante erwähnt: «[zu] jmdm/etwas Sorge haben». Ein ent-

sprechender Eintrag ist auch in der Wikipedia unter *Helvetismen* zu finden: *Sorge tragen zu jemandem/etwas* (sich kümmern um jemanden),²² wie im folgenden Beispiel:

- (64) ...fordert Pro Natura eine Jagd, die zur Natur Sorge trägt. (CHTK: Anonymus. 2000. «Naturschutz und Jagd». (S. 4))

Im deutschen VALBU ist hingegen nur eine Beschreibung ohne Objekt enthalten, und zwar als Funktionsverbgefüge «Sorge tragen, dass» mit der Bedeutung «dafür sorgen dass...».

In den grossen Korpora kommt neben der Konstruktionsvariante *Sorge tragen zu* auch die Variante *Sorge tragen für* vor, während letztere in den Therapietranskripten nicht verwendet wird (Beispiel 65). Die Konstruktion *Sorge tragen für [etwas]* wird im Duden 11 als *gehoben* bezeichnet.

- (65) Die Bodelschwingschen Anstalten in Bethel tragen Sorge für eine faire Verteilung der Bekleidung. (BRZ08/SEP.13292)²³

Im *Schweizer Textkorpus* CHTK kommt *Sorge tragen* in verschiedenen Varianten vor, was möglicherweise auf regionale Besonderheiten des Dialekts schliessen lässt. Die Belege im CHTK stammen aber teilweise aus älteren Texten, die nicht den heutigen Sprachgebrauch widerspiegeln. *Sorge tragen* ist hier am häufigsten mit der Präposition *für* vertreten (22 mal), an zweiter Stelle mit der Präposition *zu* (15 mal). Daneben kommt der Ausdruck *Sorge tragen* aber auch mit dem einfachen Dativ vor (13 mal), und auch ohne Objekt, mit und ohne nachfolgendem Nebensatz (10 mal), wie in den Beispielen (66 und 67) gezeigt.

- (66) dann würden wir unserem Körper Sorge tragen. (Bel 1994)
- (67) man wird daher Sorge tragen, sie mit einem nassen Sacke zuzudecken. (Geb 1939)

Im *Korpus deWAC* kommt der Ausdruck *Sorge tragen* 7424 mal vor, das Nomen *Sorge* hat eine Häufigkeit von über 50'000. Das Verb *tragen* steht als Kollokationskandidat von *Sorge* an dritter Stelle.²⁴ Die Konstruktion *Sorge tragen für* hat eine Häufigkeit von 1161, *Sorge tragen dafür, dass...*

²²siehe URL: <http://de.wikipedia.org/wiki/Helvetismus> (15.11.2010)

²³Dieses Beispiel stammt aus dem COSMAS-Korpus.

²⁴deWAC: Lemma *Sorge*: 89267; Kollokationskandidaten: elterlich, bereiten 2509/9.08), tragen (7494/9.07), ... machen (15045/7.8).

kommt gar 4606 mal vor. Die Häufigkeit der Konstruktion *Sorge tragen zu* konnte nicht eruiert werden.²⁵

Die Tabelle (9.15) zeigt die Konstruktion *«Sorge tragen für»* und die Unterschiede zur oben beschriebenen Konstruktion *«Sorge tragen zu»* (Tabelle 9.13).

Tabelle 9.15: Konstruktion *Sorge tragen für etwas*

Attribut	Wert	Bemerkung
canForm	<i>tragen Sorge [für etwas]</i>	
crossRef	sich kümmern um	near-synonym
ptOfSpeech	Verb	(head)
head	tragen	
entryFormation	phrase	
phraseType	set-phrase	
synFrame	505	Verb + AkkO + (PräpO)
synType	funct	Funktionsverb
semType	act	unspecified activity
definition	dafür sorgen, dass etwas geschieht (Duden 11)	
natclause	jmd trägt Sorge für etwasn	
bauplan	505: 1 - - L2 subj Human 2 tragen VVFIN L0 root n/a 3 Sorge NN L2 objn n/a 4 für APPR L3 pp n/a 5 - - L4 pn Anything	
pattern	[[Human]] tragen Sorge für [[Anything]] dafür [dass + SATZ]	
construction	[tragen Sorge zu]	
subjField	general	Diskurstyp unspezifisch
textType	unspezifisch	Diskurseinheit
register	gehoben	(Duden 11)
topic	unspezifisch	Thema unspezifisch
Funktion	Eine Aktion in die Wege leiten	
JAKOB-Code	SOR	SOR=bangen (Wollen – Intendieren)

Verwandte Konstruktionen: *Sorge geben* und *Sorge haben*

Im Korpus WIL kommt einmal die Konstruktion *jemandem Sorge geben* vor (Beispiel 68). Dies scheint ebenso ein schweizerischer Dialektausdruck zu sein wie die Konstruktion *Sorge haben* (Dialekt «Sorg ha») im Beispiel (69)

(68) Korpus WIL

01 T: also dass sich der Ärger nicht so zeigt sondern so eher

²⁵weil Kollokationen mit *Sorge tragen* und *zu* immer den *Infinitiv mit zu* berücksichtigen und dadurch das Resultat verfälschen.

- im Sinne von ,
 02 ja also eben also eben das sind ja halt so wirklich
 so schwache schwache Leute
 03 also dass dass der Ärger eher so zum Ausdruck kommt als
 nicht als Ärger
 04 sondern eben das das sind halt Schwache
 05 denen muss man Sorge geben (p:03)
 06 es hat ja dann ein bisschen + eine
 07 P: ja ja +
 08 T: Position wo Sie dann stärker + sind
 09 P: ja + ja ja ja + ja
 10 T: ja + oder ? und + das
 11 P: ja ja +
 12 T: ja
 13 P: jaa , jaa , ich würde sogar meinen
 14 es hat mit ziemlich viel Verachtung zu tun

- (69) T: also Sie haben jetzt das Zutrauen in sich selbst so hervorgehoben, man könnte auch sagen, ja, irgendwo wissen Sie es ja, ist es nicht eher auch, ein äh, ein Sorge haben ein Vorsichtig-sein mit der Beziehung mit dem Anderen ? (WIL.229.060)

Beide Dialekt-Konstruktionen können bedeutungsmässig in die Gruppe der Konstruktionen <tragen Sorge zu etwas/jemandem> eingeordnet werden, die Funktionen der Rücksichtnahme, des sorgfältigen Umgangs und des Beschützens beinhalten.

9.2.4 Konstruktionen mit «Zeug»

Unter die Konstruktionen mit Zeug fallen Verbkonstruktionen wie «ein Zeug machen», «in ein Zeug kommen» oder «in ein Zeug bringen», aber auch das Nomen «Zeug»²⁶ allein ist als Konstruktion sehr interessant, weil die Kernbedeutung dieses Nomens sehr vage oder kaum vorhanden ist, und weil an diesem Ausdruck die dynamische Entwicklung der Bedeutung im vorliegenden Kontext gut nachverfolgt werden kann. Die Konstruktionen zeigen einerseits, dass das Wort «Zeug» in einem breiten Spektrum von Konstruktionen verwendet wird, die nicht als synonym eingestuft werden können, sie zeigen aber auch, dass Konstruktionen variabel sind, einerseits auf der lexikalischen Ebene, andererseits auch im spontanen Sprachgebrauch individuelle Variationen zulassen.

²⁶Die Nomen Zeug und Zeugs kommen beide in den Transkripten vor und werden als Synonyme behandelt.

Das folgende Beispiel ist eine *Erzählung*, die Wilma in der ersten Stunde präsentiert.²⁷

(70) Erzählung «Typisch Frau» (WIL.001.249)

01 P: ich kann mich schleCHT (-)so (-) dann äussern;
 02 T: sie haben diesem freund NICht gesagt- eh du,
 03 das ist (--) das ist dann also MEIne Sache,
 04 ob ich ein Kind will oder nicht oder
 05 P: ich habe es ihm dann SPÄter eeeh::: noch gesagt;
 06 T: ja
 07 P: eh::- er habe mich dann SCHAUrig verLEtzt mit dieser bemerkung;
 08 und dann hat er dann (--) aber AUCh wieder sehr wie mein vater
 09 auch=eh:- gefunden was ICH jetzt für ein ZEUGs mache- (---)
 10 und das sei TYpisch FRAU so Zeugs immer so auseinander
 11 zu beineln und zu an(h)alys(h)ieren (h)oder(h) (-)
 12 und SCHON haben wir natürlich wieder krAch gehabt
 13 miteinander[oder?
 14 T: [ja=ja=ja

Die Klientin Wilma schildert, wie sie in einer früheren Situation Mühe hatte, gegenüber dem ehemaligen Freund ihre Enttäuschung und Empörung auszudrücken; sie verstummt in solchen Situationen und kann nicht unmittelbar reagieren. Der Therapeut greift das Thema der vergangenen Situation inhaltlich auf und fragt, ob es nicht die Wilmas Sache sei, sich dazu zu äussern (2-4). Sie sagt nun zwar, dass sie dies später nachgeholt habe, geht aber nicht weiter auf den Inhalt der Meinungsverschiedenheit ein, sondern vergleicht die Reaktion des Freundes mit einem vom Vater bekannten Gemeinplatz: Frauen veranstalten grosse Aufregungen um nichts, und Frauen wollen alles genau auseinander nehmen und analysieren, was zum offenbar wiederholten Male zum Streit führt.

Wilma räumt zwar ein, dass sie die Bemerkung über das Kinderkriegen verletzt hat, sie geht aber nicht weiter auf ihre Gefühle ein, sondern wechselt auf eine verallgemeinernde Ebene und erklärt, was Männer, wie hier der Vater und der Freund, *typischerweise* in solchen Situationen sagen. Dazu verwendet sie mehrere formelhafte Konstruktionen. *Ein Zeugs machen* bezeichnet hier, eine Sache aufzubauschen und ihr eine Wichtigkeit zu geben, die ihr nicht zukommt. Die Verbindung mit *Zeug* macht die Konstruktion besonders vage und für viele Inhalte verwendbar. Formelhaft ist auch die Wendung *typisch Frau* und die zweite Verwendung von *Zeugs*, das zerlegt

²⁷ Dieses Beispiel ist auch im Kapitel zur Erzählanalyse JAKOB (2.3) zu finden und ist nach dem GAT-Standard transkribiert (Selting et al., 1998).

und analysiert werden soll. Dies hebt die Aussage auf eine allgemeingültige Ebene und entzieht ihr gleichzeitig die Aussagekraft für die konkret vorliegende Situation.

Die Klientin Wilma erzählt im Vorfeld über eine vergangene kurze Beziehung, in der sie sich immer wieder bevormundet, manipuliert und als «Eigentum» fühlte; dies ertrage sie nicht, es sei «Terror». Als sie einmal in einer nebensächlichen Bemerkung das Thema Schwangerschaft erwähnt, zieht sich der Freund zurück, indem er ihr rät, sich gut zu überlegen, ob sie noch Kinder haben wolle. Dies ist die verletzende Bemerkung, auf die sich der Transkriptausschnitt bezieht.

In den weiteren Redezügen nach dem zitierten Gesprächsausschnitt geht der Fokus ganz vom angesprochenen Thema des Kinderkriegens und vom damit konnotierten Thema der familiären Geborgenheit weg. In den unmittelbar anschliessenden Redezügen kommt Wilma dreimal auf die Formel *typisch Frau* zurück, sie «reagiere allergisch» darauf, und sie argumentiert gegen die «unfairen Vorurteile». Der Therapeut greift darauf die letzte Bemerkung (Zeile 12) auf und fragt nach dem Streit, den die Klientin als «zwei Stunden anschweigen» charakterisiert.

Die Verwendung der formelhaften Allgemeinaussagen führt vom aktuellen persönlichen wichtigen Gesprächsthema weg. Die Klientin drückt damit aus, dass sie als Opfer der clichéhaften Attacks nichts ausrichten kann und weicht damit einer Problemkonfrontation aus. Die Konstruktionen: *ein Zeug machen* und *Zeug machen* die Aussagen vage und für jede Situation anwendbar, nehmen der Problemdarstellung die Prägnanz. *Typisch Frau* verallgemeinert die Situation und führt auf einen andern Schauplatz. Der Schlusssatz («und schon haben wir natürlich wieder Krach») führt definitiv das neue Thema ein (und vom alten weg).

Weitere Konstruktionen mit Zeug

Die Tabelle (9.16) zitiert alle Konstruktionen mit *Zeug* aus dem Korpus WIL, wie sie sowohl von der Klientin als auch vom Therapeuten verwendet werden²⁸ (siehe auch Luder, 2009).

Die aufgeführten Konstruktionen haben gemeinsam, dass sie eine Handlung bezeichnen, und dass das Tun oder Erleben der handelnden Person als «Aufregung», «Theater», «Durcheinander» beschrieben werden kann. Alle Beispiele sind Funktionsverbgefüge, d. h. die Verben sind semantisch weitgehend entleert, die Bedeutung des Verbalausdrucks entsteht durch den no-

²⁸Die Spalte *Anzahl* gibt an, wie oft die Konstruktion von der Klientin verwendet wird, in Klammer die Verwendung durch den Therapeuten.

Tabelle 9.16: Konstruktionen mit <Zeug> im Korpus WIL

Konstruktion	Paraphrase	Anz.	Kontext (Beispiele)
ein Zeug machen	Aufhebens machen, eine Aufregung veranstalten	4	...äh, ich mache da überhaupt kein so Zeugs mehr darum, oder ...
ein Zeug haben	ein Durcheinander haben	5 (1)	...und dann hatte ich ein ziemliches Zeugs äh ich raste quasi so kurz vor Feierabend dann äh so richtig los ...
ein Zeug sein, ein Zeug geben	mit expletivem <i>es</i> gebraucht: es ist ein Durcheinander, ein Theater, eine Aufregung. Es ist eine Sache, es gibt Stress.	5 (1)	...das war noch so äh, war auch immer noch so ein Zeugs mit ihm ...
im Zeug sein	im Stress sein, in Aufregung sein.	3 (1)	...also wenn ich wirklich, ganz ausgesprochen im Zeugs bin und unter Druck stehe ...
in ein Zeug bringen	in Aufregung bringen, durcheinander bringen.	2 (1)	...es ist eine wichtige Sache weil, das brachte sie ja nach- nachträglich sehr ins Zeug («is Züüg bracht») ...
in ein Zeug kommen, ins Zeug kommen	in grosse Aufregung geraten.	5 (3)	...und ich kam damals plötzlich so in ein Zeugs ...

minimalen Bestandteil. *Zeug*/*Zeugs* übernimmt als Wortmetapher die Funktion, einen schwierig und nicht eindeutig zu beschreibenden mentalen Zustand darzustellen. Der Quellbereich der Metapher bezeichnet ursprünglich einen konkreten Gegenstand, nämlich Stoff, daneben wird *Zeug* aber für viele nicht näher bestimmte Gegenstände und Sachverhalte gebraucht, oft in einem abwertenden Sinn. Der Zielbereich der Metapher ist der mentale Zustand, der verschiedene Ausprägungen annehmen kann. Die Bezeichnung dieses Zustandes mit *Zeug* erspart den Gesprächsteilnehmern einerseits die Suche nach dem treffenden Ausdruck, ermöglicht ihnen andererseits, die Bedeutung des Zustandes vage zu halten und eventuell auch herunterzuspielen (in den Wörterbucheinträgen wird auf die abwertende Funktion der Verwendung von *Zeug* hingewiesen).

Die relative Vielfalt der oben aufgeführten Konstruktionen gruppiert sich um eine *Kernbedeutung* oder um ein *Kernthema*, das mit Aufregung, Durcheinander, Stress beschrieben werden könnte. Aus dieser Kernbedeutung können verschiedene Konstruktionen und Variationen abgeleitet werden; diese sind variabel, sowohl auf der lexikalischen Ebene als auch im spontanen Sprachgebrauch.

Zeug in den Korpora

Tabelle 9.17: Konstruktionen mit «Zeug»

Konstruktion	WIL	AMA	GUS	AGD	CHTK	deWAC
Zeug	110	68	9	879	273	14'500
«ein Zeug machen»	4	(3)	0	(18)		(165)
«ein Zeug haben»	6	0	0			(2291)
«ein Zeug sein/geben»	6	0	0			
im Zeug sein	4	0	0			
in ein Zeug bringen	3	0	0			
in ein Zeug kommen,						
ins Zeug kommen	8	0	0	(4)		
am Zeug flicken	0	0	0	1	2	76

Die Tabelle (9.17)²⁹ listet das Vorkommen der verschiedenen Konstruktionen um den Ausdruck *Zeug* auf.

- WIL: Im Korpus WIL kommt das Lemma *Zeug* 110 mal vor. Wichtigste Kollokationskandidaten³⁰ sind in absteigender Reihenfolge: *wütend*, *herum*, *solch*, *ins*, *drin*, *furchtbar*, *dies*, *im*, *sonst*, *wobei*, *so*, *Sache*, *da*, *mit*, *in*, *kommen* (5), *haben* (26), *machen* (5). «Haben» kommt also oft zusammen mit «Zeug» vor, aber nur 6 mal im Sinne der Konstruktion «ein Zeug haben».
- AMA: Das Lemma *Zeug* kommt 68 mal vor. Kollokationskandidaten sind hier vor allem Adjektive: *dumm*, *solch*, *laut*, *blöd*. Verbverbindungen sind z. B. *Zeug nehmen* (Medikamente) und *Zeug finden*. Der

²⁹Zahlen in Klammern bedeuten, dass entsprechende Wortverbindungen zwar vorkommen, aber nicht in der Konstruktionsbedeutung; die Wortverbindung «Zeug machen» entspricht nicht der Konstruktion «ein Zeug machen» mit obligatorischem unbestimmtem Artikel!

³⁰Rangfolge gemäss dem *logDice*-Assoziationsmass (Rychlý, 2008). Das Zusammenhangsmass ist nicht nur von der Häufigkeit, sondern vor allem von der Exklusivität des Zusammen-Auftretens abhängig.

Ausdruck *Zeug machen* kommt bei Amalie dreimal vor, aber immer im Sinne von «Sachen erledigen», d. h. nicht in der gleichen Bedeutung wie im Korpus WIL (Schweizer-Mundart).

- GUS: Das Lemma *Zeug* kommt 9 mal vor, aber keine der Konstruktionen aus der Tabelle (9.17). Auch hier sind vor allem Adjektive mit *Zeug* assoziiert: *fachliches Zeug*, *alles mögliche Zeug* (2), Satzfragmente sind *werde ich mit meinem Zeug nicht fertig*, *so normal mein Zeug machen*, *die nehmen das Zeug so ernst*, *gehört mir das ganze Zeug*. Auch hier wird *Vagheit* als wichtige Eigenschaft dieses Ausdrucks deutlich.
- AGD: In der *Datenbank Gesprochenes Deutsch* kommt *Zeug* 879 mal vor, die Suche nach *Zeugs* bringt zusätzliche 28 Fundstellen. Die oben erwähnten Konstruktionen kommen nicht vor, aber zum Beispiel die Konstruktion «am Zeug flicken»: «...ist es aber ja Mode eben geworden der Kirche am Zeug zu flicken» (FR153).

Der Ausdruck *Zeug machen* erzielt 18 Treffer; keiner der Treffer ist aber die Konstruktion «ein Zeug machen» aus dem Korpus WIL. Beispiele: «...und dass ich sehr bald, mit sechzehn Jahren mir schon das Zeug machen musste» (ZWZ06); «Die Erna macht ihr Zeug allein, das ist ja nicht viel» (ZWJ68).

Das Gleiche gilt für die Wortverbindung *Zeug kommen*, die viermal auftritt, aber nicht im Sinne der Konstruktion «in ein Zeug kommen». Beispiele: «ich konnte ja nicht ohne Zeug nach Hause kommen» (HE071). «Zum Distel stechen und zu dem Zeug kommt man heute fast nicht mehr» (ZWB94).

- CHTK: Im Schweizer Textkorpus, das auf geschriebener Sprache beruht, finden sich 283 Fundstellen für das Lemma *Zeug*, zusätzliche 10 mit *Zeugs*. Kollokationskandidaten können in diesem Korpus nicht eruiert werden. Auch hier wird *Zeug* oft mit Adjektiven qualifiziert, wie *dummes*, *närrisches*, *wirres Zeug*.

Die Konstruktionen aus der Liste oben kommen im Schweizer Textkorpus nicht vor. Beispiele von Wortverbindungen mit *Zeug*: «Legt man sich für einen solchen Kerl ein wenig ins Zeug» (Bel 1921). «Alwin glaubte das Zeug zu einem Jünger Casals in sich zu haben» (SaT 1958). «Da wird gelogen, was das Zeug hält» (Ztg 1911). «Da hat man auch noch Zeug und Sachen mit heimbringen können» (Fässler, Hermine. 1920. «Der Aufstieg». (S. 48)). «ich bin nicht gekommen, um

dir am Zeug zu flicken» (Bel 1997). «Was denkst du! Solches Zeug versteht doch der Bub noch nicht!» (Bel 1919).

- deWAC: Im Korpus deWAC gibt es ca. 14'500 Treffer für die Wortform *Zeug*³¹. Kollokationskandidaten in absteigender Reihenfolge sind *wirr*, *dumm*, *mächtig*, *flicken*, *legen*, *ins*, *irgendwelches*, *unnütz*, *belanglos*, *halten* («was das Zeug hält»), *Klassiker* («hat das Zeug zum Klassiker»), *Frühstückseier* (Fehler, weil 33 mal gleicher Satz im Korpus!), *unverständlich*, *allerlei*, *reden*, *schmecken*, *so'n*, *stinken*, *stammeln*, *klebrig*, *labern*, *schlucken*, *sinnlos*, *schmieren*, *eklig*, *neumodisch*, *komisch*, *faseln*, *nutzlos*, *wertlos*, *schwätzen*, *albern*, *blöd*, *schmeissen*.

Die Wortverbindung *Zeug machen* kommt 165 vor, nicht aber die Konstruktion «ein Zeug machen».³²

Die Konstruktionen aus der Tabelle (9.16) kommen praktisch nur im Korpus WIL vor, einem Korpus mit schweizerdeutscher gesprochener Mundart. Es müssten weitere solche Korpora zur Verfügung stehen, um die Frage beantworten zu können, ob die Zeug-Konstruktionen aus dem Korpus WIL allgemein verbreitet sind oder ob diese Konstruktionen kreative Neuschöpfungen aus dem Kontext der Gesprächssituation sind. Dass diese Frage auch von Muttersprachlern nicht so klar intuitiv beurteilt werden kann, zeigt das Potenzial der *dynamischen Bedeutungsherstellung* beim Meinen und Verstehen und des damit verbundenen Konzeptes der *Vagheit* (Linell, 2009, S. 226).

Sind die *Zeug-Konstruktionen* kennzeichnend für den Schweizer Dialekt? In den einschlägigen Wörterbüchern ist dazu nichts zu finden; eine informelle Google-Suche speziell nach der Konstruktion «ein Zeug machen» (10.11.2008) ergab ca. 115 Treffer, eingeschränkt auf CH-Domänen nur 33 Treffer; eine Auswahl davon ist in der Tabelle (9.18) aufgelistet.³³

Belege für die Konstruktion «ein Zeug machen» in der Bedeutung von «eine Aufregung produzieren» wurden nur in den Domänen CH und AT gefunden, die Seiten aus der Domäne DE belegen diese Konstruktion nicht. Auch die grossen Wörterbücher aus dem hochdeutschen Sprachraum führen diesen Ausdruck nicht auf (Duden 10, 2002; Duden 11, 2008; Der kleine

³¹Ein Problem besteht bei der Abfrage im deWAC: Die Suche nach dem Lemma «Zeug» bringt auch das Wort «Zeuge» im Ergebnis, insgesamt 35'000 Treffer.

³²Vorkommen: Phrase *ein Zeug*: 120 mal, Kombination *ein Zeug* + *machen*: 5 mal, aber in diesen 5 Treffern besteht keine direkte Verbindung von *Zeug* und *machen*.

³³Erste Spalte: Beispielsatz, zweite Spalte: entspricht die Wortverbindung der Konstruktion «ein Zeug machen»? Dritte Spalte: Region (Internet-Domäne DE oder CH).

Tabelle 9.18: Konstruktionen mit ‹Zeug› im Internet (Luder, 2009)

Kontext	Idiom	Region
womit man kostenlos music mixen und scratchen und so ein Zeug machen kann.	nein	DE
aber man kann damit Kurvendiskussionen und so ein Zeug machen	nein	DE
werbung, die man für so ein zeug machen darf	nein	DE
Und über so ein Zeug machen die eine Prüfung.	nein	DE
Wenn ich Demos und so ein Zeug mache, benutze ich das ziemlich oft	nein	DE
Peelings und so ein Zeug mache ich nicht.	nein	DE
Da ich beruflich so ein Zeug mache, muss ich die Sendung nicht sehen...	nein	DE
wie kannst du nach deinem wochenend thread jetzt so ein zeug machen	nein	DE
Dass die Liberalen so ein Zeug machen versteht man.	ja	CH
der gesagt hatte, man solle nicht so ein Zeug machen um die 275 000	ja	CH
WiE KaNn mAn nUr sO EiN ZeUg mAcHeN Voll krass	ja	CH
man solle doch nicht solch ein Zeug machen wegen 8000 Fränkli Busse.	ja	CH
Blödsinn, diese elektronische Kunst, was die da für ein Zeug machen	ja	AT

Wahrig, 2007), und auch im DWDS (Digitales Wörterbuch der Deutschen Sprache) ist der Ausdruck nicht belegt.

Die Konstruktion *ein Theater machen* (nicht im Sinn von Theater spielen!) ist mit der Konstruktion *ein Zeug machen* verwandt und kann als Beinahe-Synonym bezeichnet werden. Sie kommt im Korpus WIL 14 mal vor, im Korpus AMA fünfmal; Gustav verwendet die Konstruktion nicht.

Weitere Funktionen

Der Ausdruck *Zeug* und Konstruktionen mit *Zeug* werden – mindestens in Therapiegesprächen – sehr häufig verwendet. Vor allem die vielfältige Verwendung des Nomens «Zeug» trägt zu diesem Befund bei. Die Konstruktion *ein Zeug machen* scheint eine regionale Besonderheit der schweizerischen/alemannischen Sprachregion zu sein, ist aber ein allgemein verwendeter Ausdruck und nicht eine individuelle Konstruktion von Wilma.

Ein Phänomen, das an den Konstruktionen mit *Zeug* sichtbar gemacht

Tabelle 9.19: Konstruktion: *ein Zeug machen*

Attribut	Wert	Bemerkung
canForm	<i>machen ein Zeug</i>	
orthVariant	machen ein Zeugs	
crossRef	machen ein Theater, machen ein Drama	
ptOfSpeech	verb	(Wortart des Heads)
head	machen	Funktionsverb als Head
entryFormation	multi word/phrase	
phraseType	Idiom	
synFrame	500	Satzmuster nach Wahrig
synType	funct, trans	function verb, transitiv
semType	act	unspecified activity
definition	sich kompliziert verhalten, viel Aufhebens machen	
natclause	jmd macht ein Zeug	
bauplan	500: 1 - - L2 subj Human 2 machen VVFIN L0 root n/a 3 ein ART L4 det n/a 4 Zeug NN L2 obja n/a	
pattern	[[Human]] machen {ein Zeug ein Zeugs}	
construction	[ein Zeug machen]	
subjField	Therapiegespräch	Diskurstyp
textType	linguistischer Marker	Diskurseinheit
register	umg, CH	umgangssprachlich, Schweizer Dialekt
topic	lokales Diskursthema	lokales Pattern
Funktion	Ein Verhalten als Übertreibung darstellen	
JAKOB-Code	DAR-ERR	DAR=Darstellen (Handeln – Präsentieren); ERR=Erregen (Geschehen – Vorgehen)

werden kann, ist die *Emergenz* der Bedeutung. Im alltäglichen Sprachgebrauch wird auf Bekanntes zurückgegriffen und gleichzeitig der zukünftige Sprachgebrauch entwickelt und vorgespart; dies gilt sowohl für die rezeptive wie für die produktive Ebene.³⁴ Im Verlauf der Therapie von Wilma werden der vage Begriff *Zeug* und die *Zeug-Konstruktionen* zunehmend mit einem Bedeutungspotenzial versehen, das immer noch vage ist, das aber für die beiden Gesprächsteilnehmer eine Übereinstimmung in der Vagheit ermöglicht, das sich im Formelhaften niederschlägt. Für diese Annahme spricht auch, dass der Therapeut im Laufe der Zeit die Zeug-Konstruktionen, die initial von Wilma ins Gespräch gebracht wurden, übernimmt und selber

³⁴Vgl. das *lexical priming* (Hoey, 2005). Obschon die beiden Begriffe aus verschiedenen wissenschaftlichen Diskursen stammen, bezeichnen sie doch ein vergleichbares Phänomen.

verwendet. *Zeug* wird im Laufe der Gespräche zu einer Formel, die von Patient und Therapeut gemeinsam aufgebaut wird.

9.2.5 Emotionskonstruktionen mit «verjagen», «explodieren» und «Dampf»

Ausgangspunkt für die folgenden Konstruktionen ist das Verb «verjagen»; Wilma verwendet wiederholt die Ausdrucksweise «...dann hat es mich verjagt...». Das Verb *verjagen* in dem hier verwendeten Sinn ist spezifisch für den Schweizer Dialekt und kann als Beinahe-Synonym zum Verb *explodieren* betrachtet werden. «Es verjagt mich» heisst soviel wie «ich explodiere»; man kann allerdings nicht sagen «ich verjage». Wilma verwendet beide Ausdrucksweisen, verjagen (19 mal) etwas weniger als explodieren (33 mal).

(71) Wilma: «dann hat es mich verjagt» (WIL.004.387)

01 P: und ich war sauber und glatt um viertel nach eins in §Vorstadt
 02 und die sind nicht mehr zu Hause. und äh, waren also schon gegangen
 03 T: hmhm
 04 P: und dann äh, habe ich
 05 dann äh hat es mich verjagt ,
 06 ich musste erst weinen, oder?
 07 und bin dann äh im Bus wieder zurück nach §Grossstadt, oder?

Die Klientin Wilma kommt pünktlich am Wohnort der Eltern an, diese sind aber schon nicht mehr zu Hause, worauf Wilma ihre Gefühle nicht mehr unter Kontrolle hat und weint. Schliesslich fährt sie mit dem Bus wieder zurück.

Wilma erzählt die Szene, um zu zeigen, wie sich die Mutter gegenüber dem Vater und auch ihr durchsetzte («so eine Situation wo sie wieder mal über das Leben bestimmte»), und warum sie sich vor mehreren Jahren von ihr distanzieren musste. Wilma sollte zu einer abgemachten Zeit bei ihren Eltern sein, um Verwandte zu besuchen, war knapp an der Zeit und meldete telefonisch der Mutter, dass sie den vereinbarten Zeitpunkt nur ganz knapp schaffen werde. Als sie doch noch rechtzeitig ankommt, stellt sie fest, dass die Eltern bereits abgefahren sind. Im Nachhinein stellt sich heraus, dass die Mutter dem Vater ausgerichtet hatte, dass Wilma möglicherweise gar nicht komme und sie so übergangen hat.

Auch im zweiten Gesprächsausschnitt erzählt Wilma von einer Situation, in der sie die Kontrolle über ihre Gefühle kurzzeitig verloren hatte und bezeichnet dies als einen «saublöden Moment» (Beispiel 72). Während es im ersten Beispiel Gefühle der Trauer sind, die sich eruptiv geäussert haben, sind es im zweiten Beispiel eher Gefühle der Wut.

(72) Wilma: «es hat mich verjagt» (WIL.315.068)

01 P: und der zweite Fall
02 da musste ich mich echt entschuldigen
03 da ist er mir wirklich , in einem absolut ,
saublöden Moment begegnet
04 wo es mich nämlich sowieso gerade verjagt hat , oder ?
05 und das war es?
06 das war es
07 und das ganze riesige Theater, das war hin.
08 wobei äh der, Abteilungsleiter hat dann nachher gesagt
09 er hätte eigentlich so-;
10 er hätte auch einen Fehler gemacht

Im Korpus Amalie kommt das Verb *verjagen* nicht vor; das Verb *explodieren* verwendet sie 11 mal als Beschreibung für einen Wutausbruch, einmal im «wörtlichen» Sinne («ich dachte schon bei mir explodiert die ganze Küche»).

(73) Amalie: «zweimal explodiert» (AMA.98.4278)

01 P: sie hat abgeschrieben
02 ich war ganz ruhig geblieben
03 sie war explodiert und eh
04 ich hab das bloß eben vor der Klasse gemacht
05 und sie war dann gekommen und hat vor der ganzen Klasse
mich angeschrien
06 in der Schule explodier ich eigentlich ganz , ganz arg selten
07 entweder weil ich mir vorher die Sache mir überlegt hab
08 oder weil ich Schreien wahnsinnig dumm finde
09 ich hab mal ' nen Lehrer gehabt
10 der schrie und
11 und da weiss nicht
12 da hat sich immer so alles bei dem verändert
13 er kriegte so ' nen komischen Kopf
14 und das fand ich so lächerlich
15 und das war ein ganz schlimmes Beispiel
16 ich schrei höchstens zu Hause , ganz kurz , aber dann sehr explosiv
17 (Pause) . in den Ferien bin ich zweimal explodiert
18 doch , das kann ich schon
19 nein , ich glaub eher
20 da ist , das Schämen ist drin
21 denn ich hab mich letztesmal doch ziemlich geschämt
22 wenn ich so zurückdenke
23 als das mit dem Träumen da war
24 war mir glaub ich doch ziemlich peinlich . und eben
25 weil Sie da an dem Tisch mit sassen , nicht
26 so eher seh ich das
27 dass ich mich dann eben für meine Mutter geschämt hab

28 oder für mich selber
29 das ist in dem Fall eins

Der Gesprächsausschnitt (74) zeigt, dass die Konstruktion «explodieren» nicht nur für die Beschreibung von Wutausbrüchen zur Anwendung kommt; man kann auch explodieren vor Lachen, wie dies in der folgenden Szene beschrieben wird. Explodieren vor Lachen betont hier, dass das Lachen quasi als Naturgewalt über einen hereinbricht, dass man sich nicht dagegen wehren kann.

(74) Wilma: «explodiert vor Lachen» (WIL.135.462)

01 P: oder ? und das ist natürlich so äh die Nervosität
02 das ist nochmal
03 wie soll ich sagen ?
 noch noch eine ziemliche Hemmschwelle
04 T: hmhm
05 P: und äh , aber ich kann natürlich dann absolut wirklich
 im kurligsten Moment lachen (lacht)
07 T: hmhm
08 P: nein ich darf es mir nicht vorstellen (lacht)
09 stellen Sie sich die Situation vor
10 Sie gehen das erste Mal zu einer Frau nach Hause (lacht)
11 plötzlich explodiert die vor Lachen
12 weil sie nämlich einen Frosch im Kopf hat
 (lacht wirklich , T. leise mit)
13 das ist doch wirklich (hustet , lacht)
14 da kann ich ja dem nicht sagen
15 hör mal würdest du den Frosch (lacht) ///
16 T: äh (lachen)
17 P: ich knalle dich jetzt an die Wand hoch (lacht) also nein
18 T: hmhm
19 P: oh Gott (lacht) das ist eigentlich noch eine gute Möglichkeit
 (lacht) gegen Nervosität (lacht)

Weitere Emotionskonstruktionen mit der Behältermetapher

Für diesen Bereich der explosionsartigen Entladung von Emotionen gibt es eine ganze Reihe von verschiedenen Konstruktionen, die sich fast alle auf die konzeptuelle Metapher ÄRGER IST ERHITZTE FLÜSSIGKEIT IN EINEM BEHÄLTER beziehen (Tabelle 9.20, vgl. Kapitel 4.3.2).³⁵

Die erwähnte konzeptuelle Metapher müsste allerdings auf Grund der Korpusevidenzen abgeändert werden zu GESTAUTE EMOTION IST ERHITZTE FLÜSSIGKEIT IN EINEM BEHÄLTER.

³⁵Die Doppelzahlen dokumentieren das Vorkommen der metaphorischen Konstruktion gegenüber dem Vorkommen der «wörtlichen» Bedeutung.

Tabelle 9.20: Konstruktionen von «explodieren» bis «Dampf ablassen»

Konstruktion	WIL	AMA	GUS	AGD	CHTK	deWAC
verjagen	19/2	0	0	0/8	0/42	?/2962
explodieren	33	12	0	1/7	ca.20/43	?/11810
Kragen platzen	0	2	1	0	2	712
Dampf	9	6	4	58	308	7264
Dampf ablassen	6	5	3	1	3	412
Dampf machen	2	0	0	0	0	578
seinen Dampf loswerden	1	0	0			
Dampf aufladen	1	0	0			
Überdruck ablassen	1	0	0			
Frust ablassen	1	0	0			

- Es zeichnen sich drei *Bedeutungsvarianten* von *verjagen* ab: (1) <es verjagen jmdn>: Es verjagt eine Person: diese Person verliert die Geduld, sie macht ihrem Ärger Luft. Wilma braucht diese Konstruktion (wie in den Beispielen beschrieben) 19 mal. Zweimal braucht sie daneben die wörtliche Bedeutung *jmdn [von einem Platz] verjagen* (2). Daneben gibt es eine weitere Bedeutung, z. B. im Satz «Es hat den Benzintank verjagt» (3). Letztere stellt den Quellbereich der metaphorischen Bedeutung (1) dar. Konstruktion (1) und (3) repräsentieren umgangssprachlichen Schweizer Dialekt. Im Korpus AGD, aber auch im Korpus CHTK gibt es keine Konstruktion im Sinne von (1).
- Zu *verjagen* gibt es gemäss Korpus WIL mehrere Variationen: So braucht sie den Ausdruck *fast verjagen* («es hat mich fast verjagt»). *Verjagen* kann auch in Bezug auf die Emotion präzisiert werden: «Es hat mich fast verjagt vor Wut». Ausserdem kommt bei Wilma die Kombination mit der Konstruktion <kurz spitz> vor («da hat's mich kurz spitz verjagt»).
- Zwei weitere umgangssprachliche Variationen im Korpus WIL (aber vom Therapeuten gesagt) sind die schweizerdeutschen Ausdrücke *den Nuggi herausjagen* («es hätt ihre de Nuggi usegagt») und *den Deckel abjagen* («jetzt jagts mer dä Deckäl ab»). Und später wandelt Wilma diese Wortverbindung ab zu («denn häts mer no ganz dä Nuggi glupft»).³⁶ Diese Redensart ist eher in der Variante «es hat mir den

³⁶Schweizerdeutsch: lüpfen oder lupfen = heben.

Deckel gelupft» verbreitet.

- *explodieren* kann als Beinahe-Synonym zu *verjagen* betrachtet werden. Beide Emotionsausdrücke können sich sowohl auf die Emotionen Wut, Freude oder Trauer beziehen.³⁷
- <platzen der Kragen jmdm>: Im Korpus WIL kommen weitere Konstruktionen mit *platzen* vor: <platzen vor Lachen>, platzen vor Stolz, <Sicherung platzen jmdm> («und da ist mir dann die Sicherung geplatzt»), <platzen wie eine Seifenblase>.

Nachdem die Konstruktion <es verjagen jmdn> in keinem der Korpora ausser bei Wilma dokumentiert ist und auch im Schweizer Textkorpus und im Schweizer Wörterbuch (Meyer & Bickel, 2006) keine entsprechenden Hinweise gefunden wurden, ergab ein Google-Suche im Internet folgende Treffer (30.11.2010):

Aber wir haben das Kind mit dem Bade ausgeschüttet und vergessen, dass wir Spielregeln brauchen, um ein sinnvolles Zusammenleben zu regeln. Auch wenn es manchmal unangenehm ist, diese durchzusetzen.

(Interviewer: Das ist der Hauptgrund für den Lehrermangel?)
Es ist ein wesentlicher Teil der Problematik, der selten beachtet wird. Ich kenne viele Lehrer, die es fast «verjagt», weil sie damit so grosse Probleme haben. Da sage ich: Entweder kannst du es, oder du musst damit aufhören. (Interview mit Bildungsexperte Peter Schwarzenbach in der BAZ, 3.7.2010)

Das Zitat stammt aus der Basler-Zeitung, die Konstruktion wird in Anführungszeichen gesetzt, offenbar um das Mundartliche hervorzuheben. Das zweite Zitat stammt aus einem Forumsbeitrag:

Als ich heute nach hause kam, fand ich das zimmer mit dem juwel-aq drin dunkel vor, es hat die sicherung rausgejagt. als ich dann genauer schaute, stellte ich fest, dass es den heizstab in der mitte verjagt hat. temperatur ist 25°C, die fische leben bis jetzt noch alle. nun frage ich mich: ist die heizflüssigkeit giftig für die fische? (Forumsbeitrag auf der Internetseite www.aquarium.ch, 15.11.2010)

³⁷ Leider wurde bei der *Transkription* der Wilma-Gespräche in 15 dokumentierten Fällen *verjagen* durch *explodieren* ersetzt. Ob weitere der mit dem Verb *explodieren* transkribierten Stellen eigentlich das Verb *verjagen* beinhalten, müsste anhand der Tonaufnahmen überprüft werden.

Die Tabelle (9.21) zeigt die Konstruktion <verjagen>:

Tabelle 9.21: Konstruktion: <verjagen>		
Attribut	Wert	Bemerkung
canForm	<i>verjagen</i> [es] [jmdn]	
crossRef	explodieren, platzen der Kragen	
ptOfSpeech	verb	
entryFormation	single word	
phraseType	idiom	
synFrame	501	Subjekt <es> + AkkO
synType	main-vb / mid	Hauptverb / Mittelverb
semType	achiev	achievement – plötzlicher Zustandswechsel
definition	es hat mich verjagt, ich habe die Geduld verloren	
natclause	es verjagt jmdn	
bauplan	501: 1 es PPER L2 subj n/a 2 verjagen VVFIN L0 root n/a 3 - - L2 obja Human	
pattern	{es} verjagen [[Human]] ({vor [[Psych]])}	
construction	es ([MOD]) verjagen (vor [N])	
subjField	Therapiegespräch	Diskurstyp
textType	Narrativ	Diskurseinheit
register	umgangssprachlich, Dia- lekt	lokale Diskursmuster
topic	Gefühlsausbruch	lokales Pattern
Funktion	Emotion heftig zum Ausdruck bringen	
JAKOB-Code	SPA-DES	SPA=Anspannen (Geschehen – Vorgehen), DES=Zerstören (Schaffen – Ordnen)

Fragen zur Lexikonrepräsentation der Konstruktion verjagen Eine Frage zum *Lexikonaufbau* betrifft die beiden Konstruktionen *verjagen* und *explodieren* gleichermaßen: Wie wird das Verhältnis der beiden Bedeutungsvarianten im Lexikon abgebildet? Soll es je einen Lexikoneintrag geben für die *wörtliche Variante* von *verjagen* und *explodieren* («es verjagt den Benzintank», «der Benzintank explodiert») und einen separaten Eintrag für das Form-Bedeutungspaar *explodieren* resp. *verjagen* im Sinne von *einen Gefühlsausbruch haben*? Die Frage bleibt bei dieser Variante, wie die Verbindung zwischen den beiden Konzepten (literal und figurativ) geschieht. Nach den Richtlinien für Verbpattern wäre es möglich, die wörtliche und die figurative Variante im gleichen Eintrag (unter dem gleichen Verbpattern) abzubilden:

(75) [[Human | Container | Physical Entity]] explodieren [NO OBJ]

Diese Variante hätte den Vorteil, dass das wörtliche und figurative Konzept aufeinander bezogen bleiben, denn im mentalen Lexikon müssen die beiden Bedeutungsvarianten für die Metaphern- und Metonymiebildung nahe beieinander liegen.

Eine Konstruktion, die sich ebenfalls auf die Behältermetapher bezieht, ist der metaphorische Ausdruck *Dampf ablassen*, der in allen drei Therapiekorpora verwendet wird. Es folgt ein Beispiel aus dem Korpus WIL:

(76) Wilma: Dampf ablassen (WIL.263.087)

01 P: hmhm oder äh wenn ich ins Bett gehe
 02 dieses kommt mir noch in den Sinn + jenes
 03 T: hmhm +
 04 P: und dann äh immer meine Autosuggestion damit aufhören
 ist anstrengend
 05 ich habe am Morgen im Moment Mühe mit aufstehen
 06 und ich denke es ist deswegen
 07 T: ja klar
 08 P: weil etwa das Limit erreicht ist
 09 T: ja , hmhm
 10 P: aber eben ich kann ja nicht viel ändern
 ausser brüllen eben so Dampf ablassen

Wilma beschreibt eine geschäftliche Situation, in der sie von den Ansprüchen ihrer Geschäftskollegen überfordert ist und sich zur Wehr setzt. Sie reagiert mit «Brüllen» und «Bellen», wie sie sich ausdrückt, und kann so *Dampf ablassen*.

(77) Amalie: Dampf ablassen (AMA.111.186)

01 P: ...ich weiss nicht
 02 ich ich weiss wirklich nicht =
 03 warum ich das nicht endlich mal fertig bring
 04 da ein bisschen Dampf abzulassen und zu sagen
 05 okay ich weiss doch wie er ist

Amalie erzählt von ihren Besuchen zu Hause und wie sie sich immer furchtbar darüber aufregt, dass der Vater ihr beim Autowaschen helfen will. Sie kennt aber diese Situationen seit längerer Zeit und fragt sich selber, wieso sie nicht in der Lage ist, sich zu besinnen und ihre Emotionen zu kontrollieren.

(78) Gustav: Dampf ablassen (GUS.180.024)

01 T: im Gegenteil . das ist so ein Punkt ,
 02 wo der Ärger sitzen könnte .
 03 den Sie mir gegenüber immer wieder und immer wieder erleben .
 04 wo es nicht darum geht ,
 05 dass Sie hier Dampf ablassen ,
 06 sondern wo die ganz konkrete Beziehung Sie ärgert .
 07 dass es Erwartungen auf Ihrer Seite gibt ,
 08 denen ich nicht (12) nachkomme .

Im Beispiel (78) bezieht sich der Therapeut auf die Situationen, in denen der Klient (Gustav) in der Therapiestunde seine Wut und seinen Ärger zur Sprache bringt; Emotionen, die sich offenbar aufgestaut haben und die er in die Therapiesitzung mitbringt.

In allen drei Beispielen wird die Metapher «Dampf ablassen» mit der Bedeutung verwendet, die in Wörterbüchern beschrieben wird, wie z. B. im Duden 11: «seinen Ärger, seine Wut abreagieren. ...damit der Druck sich verringert und der Kessel nicht explodiert». Die Tabelle (9.22) zeigt die Konstruktion <Dampf ablassen>.

In der IVP-Liste³⁸ (Söhn, 2006) wird der Phraseologismus *Dampf ablassen* dokumentiert, siehe (7.2.5). Spezielle, in der obigen Konstruktionsbeschreibung nicht enthaltene Merkmale betreffen:

- Metaphorisch: ja
- Dekomponierbar? ja;³⁹ Dampf-ÄRGER, ablassen-ABBAUEN
- Valenz: intern: objNP; extern: subjNP
- passivierbar: ja (aber oft in wörtlicher Lesart)
- Nicht-Negation? ja;

Norms und Exploitations

Neben den Konstruktionen *Dampf ablassen* und *Dampf machen* kommen im Korpus WIL weitere Wortverbindungen mit *Dampf* vor, einmal die Variante «der Chef will seinen Dampf loswerden» und einmal seitens des Therapeuten die Frage «...wenn sie sagen Dampf ablassen, wo laden sie dann auf? Dampf?». Ausserdem kommen je einmal die Konstruktionen *Überdruck ablassen* und *Frust ablassen* vor. Die Vorkommen dieser Konstruktionen in den Korpora sind in Tabelle (9.23) dargestellt (vgl. Luder, 2010).

³⁸IVP: Idiomatisierte verbale Phraseologismen.

³⁹gemäss Söhn vollidiomatisch, aber dekomponierbar, d. h. kompositional, da «man sieht, in welcher Weise die Bestandteile zur Gesamtbedeutung beitragen» (Söhn, 2006, S. 40).

Tabelle 9.22: Konstruktion: *Dampf ablassen*

Attribut	Wert	Bemerkung
canForm	<i>ablassen Dampf</i>	
crossRef	seinem Ärger Luft machen	near-synonym
ptOfSpeech	verb	(Wortart des Heads)
head	ablassen	
entryFormation	phr	
phraseType	Metapher	
synFrame	500	Verb + AkkO
synType	main-vb, trans	
semType	achiev	achievement – plötzlicher Zustandswechsel
definition	seinem Ärger Luft machen, seinen Ärger, seine Wut abreagieren (Duden 11)	
natclause	weil jmd Dampf ablässt	
bauplan	500: 1 - L3 subj Human 2 Dampf NN L3 obja n/a 3 ablassen VVFIN L0 root n/a	
pattern	[[Human]] ablassen {Dampf}	
construction	[Dampf ablassen]	
subjField	Therapiesgespräch	Diskurstyp
textType	unspezifisch	Diskurseinheit
register	umgangssprachlich	lokale Diskursmuster
topic	lokales Diskursthema	lokales Pattern
Funktion	Emotion zum Ausdruck bringen	
JAKOB-Code	FRE-ENT	FRE=Befreien (Schaffen – Binden), ENT=Entspannen (Geschehen – Vorgehen)

Auf dem Hintergrund der Theorie der *norms and exploitations* (Hanks, 2004; Hanks & Pustejovsky, 2005; Hanks, in press) und auf Grund der vorliegenden Häufigkeitsangaben aus den verschiedenen Korpora können die mit *E/* bezeichneten Konstruktionen als *exploitations* bezeichnet werden, als Konstruktionen, die in der Interaktion und Kommunikation *dynamisch* entstanden sind und die nicht im Lexikon abgebildet sind (vgl. Kapitel 6.11).

9.2.6 Emotionskonstruktionen mit «Angst»

Ausgangspunkt für diese Konstruktion ist der Ausdruck *Angst haben*. Konstruktionen, die diesen Ausdruck beinhalten, kommen im Korpus WIL 144 mal, im Korpus AMA 228 mal und Korpus GUS 79 mal vor. Auffällig ist,

Tabelle 9.23: Norms und Exploitations: Konstruktionen mit «Dampf»

Konstruktion	WIL	AMA	GUS	AGD	CHTK	deWAC
Dampf	9	6	4	58	308	7264
Dampf ablassen	6	5	3	1	3	412
Dampf machen	2	0	0	0	0	578
<i>E/</i> seinen Dampf loswerden	1	0	0	0	0	0
<i>E/</i> Dampf aufladen	1	0	0	0	0	0
<i>E/</i> Überdruck ablassen	1	0	0	0	0	3
Frust ablassen	1	0	0	0	0	114

dass Amalie den Ausdruck häufiger braucht als Wilma, obschon das Gesprächskorpus AMA nur halb so gross ist wie das von Wilma. *Angst haben* ist ein Beispiel für eine semantisch *transparente* Konstruktion und ausserdem Beispiel eines *Funktionsverbgefüges*.

Man kann in den drei Korpora die folgenden Variationen der Grundkonstruktion *Angst haben* unterscheiden (vgl. Luder & Clematide, 2010): *Angst haben* ohne weitere Modifikation und die Variante *keine Angst haben* mit *Negation* (Bsp. 79).

(79) Gustav: Lehrproben (GUS.54.68)

01 P: ... so bei Lehrproben und solchen Dingen
 02 dass ich dann
 03 dass ich dann eben wirklich
 04 wenn ich dann drinstehe
 05 wirklich selber das Gefühl habe
 06 dass ich ruhig bin
 07 und dass ich also auch keine Angst mehr habe
 08 und nach außen hin muss das sogar noch stärker wirken
 09 ich meine...

(80) Wilma: Angst (WIL.170.157)

01 P: vor Ruinen oder so hatte ich als Kind auch Angst

(81) Amalie: Angst vor ihnen (AMA.303.148)

01 P: ja , ja , ich weiss auch nicht
 02 ich hatte plötzlich die Idee
 03 Sie würden sich an allem stossen
 04 was ich da schreibe und sagen
 05 falsch dargestellt

06 das war aber ja schon
 07 bevor Sie das mit dem Darstellen sagten
 08 T: mh
 09 P: nicht , eh , hat vielleicht auch damit zu tun
 10 dass ich eben irgendwo Angst vor Ihnen hab, ganz generell Angst
 11 T: mh
 12 P: manchmal , beim Sitzen hab ich das so gemerkt
 13 dass , da hat ich auch das Gefühl
 14 Sie nehmen mir das Gesicht auseinander

Konstruktionen in den Korpora

Tabelle 9.24: Konstruktionen mit <Angst haben>

Konstruktion	WIL	AMA	GUS	AGD	CHTK	deWAC
Angst	413	419	79	603	2495	203'129
Angst haben	169	245	32	159	334	54'602
(keine) Angst haben	14	13	1			
(grosse) Angst haben	15	5	0			
(furchtbar) Angst haben	15	5	0			
Angst haben vor	37	50	5			
Angst haben um	8	1	0			
Angst haben + Satz	15	5	0			

Die Grundkonstruktion kann fakultativ durch ein Adjektiv oder ein Adverb erweitert werden, z. B. (*grosse*) *Angst haben* oder (*furchtbar*) *Angst haben*⁴⁰. Diese Konstruktion ist in der Tabelle (9.25) dargestellt.

Eine andere Konstruktion liegt vor, wenn *Angst haben* mit einer Präpositionalphrase erweitert wird. Zwei Möglichkeiten kommen in den Korpora vor und werden auch in den Wörterbüchern aufgeführt. Mit der Präposition «vor» ergibt sich die Bedeutung «etwas fürchten» (Bsp. 82). Man kann Angst haben vor Lebewesen, Gegenständen, aber auch vor Sachverhalten. *Angst haben vor* kommt in allen drei Therapiekorpora vor (WIL: 37, AMA:50, GUS:5). Diese Konstruktion ist in der Tabelle (9.26) beschrieben.

Eine unterschiedliche Bedeutung ist jedoch bei der Ergänzung mit der Präposition «um» (*Angst haben um*) zu beobachten. Auch hier könnte das Objekt der Angst unterschiedlichen Kategorien angehören, in den 3 Grundkorpora ist es aber immer eine Angst um eine Person (Bsp. 83). Die Konstruktion kommt bei Wilma 8 mal vor, aber ausschliesslich aus dem Mund des Therapeuten. Die Bedeutung der Wortverbindung kann hier mit «sich

⁴⁰Im Korpus WIL wird 26 mal der Ausdruck «Schiss» statt «Angst» verwendet; beim Transkribieren wurde Schiss durch Angst ersetzt.

Tabelle 9.25: Konstruktion ‹Angst haben›

Attribut	Wert	Bemerkung
canForm	<i>haben Angst</i>	
crossRef	sich fürchten, haben Schiss	near-synonym
ptOfSpeech	verb	
head	haben	
entryFormation	phrase	
phraseType	set-phrase	
synFrame	500	Satzmuster zu ‹haben»!
synType	funct	Funktionsverb
semType	emotion	
definition	grosse Sorge, Unruhe empfinden	
natclause	weil jmd Angst hat	
bauplan	500: 1 - - L3 subj Human 2 Angst[sg] NN L3 obja n/a 3 haben VAFIN L0 root n/a	
pattern	[[Human]] haben ([ADV]) {[([ADJ]) Angst}	
construction	[Angst haben]	
subjField	Therapiesgespräch	Diskurstyp
textType	Narrativ	Diskurseinheit
register	unspezifisch	lokale Diskursmuster
topic	gefühle, befinden	lokales Pattern
Funktion	Kommunikative Funktion: die Emotion Angst wird ausgedrückt, Angst ohne Objekt	
JAKOB-Code	EMP-FUR	EMP=Empfangen (Handeln – Interagieren), FUR=Fürchten (Fühlen – Distanzieren)

sorgen um jmdn» umschrieben werden. Bei Amalie andererseits kommt diese Konstruktion nur einmal vor, als ‹Angst haben um den Führerschein». Die Tabelle (9.27) zeigt diese Konstruktion.

- (82) Herrgott nochmal man kann einfach zu viel *Angst haben vor Sachen*, nicht? (WIL.204)
- (83) T: muss man keine Angst haben ? muss ich *keine Angst haben um Sie*, da ?
P : nein also das ist für mich ! kein Thema. (WIL.032)

Alle diese Varianten setzen gemäss Korpusvorkommen (Therapietranskripte) voraus, dass das handelnde Subjekt des Satzes ein menschlicher Akteur ist, oder allenfalls metonymisch eine Institution.

Tabelle 9.26: Konstruktion <Angst haben vor>

Attribut	Wert	Bemerkung
canForm	<i>haben Angst vor</i>	
crossRef	fürchten, sich fürchten	near-synonym
entryFormation	phrase	
phraseType	set-phrase	
synFrame	550	Verb + AkkO + PräpO
synType	funct	Funktionsverb
semType	emotion	
definition	Angst verspüren vor etwas, vor jdm., etwas fürchten	
natclause	jmd hat Angst vor etwasm	
bauplan	550: 1 - - L5 subj Human 2 Angst[sg] NN L5 obja n/a 3 vor APPR L5 pp - 4 - - L3 pn Anything 5 haben VA L0 root n/a	
pattern	[[Human]] haben Angst vor [[Anything]]	
construction	[Angst haben vor]	
subjField	Therapiesgespräch	Diskurstyp
textType	Narrativ	Diskurseinheit
register	unspezifisch	lokale Diskursmuster
topic	gefühle, befinden	lokales Pattern
Funktion	Kommunikative Funktion: Angst benennen	
JAKOB-Code	FUR	FUR=Fürchten (Fühlen – Distanzieren)

Schliesslich gibt es die Variation *Angst haben + Satzergänzung* (z. B. dass-Satz, Infinitiv mit zu, Konjunktivsatz), wie in den Beispielen (84 und 85).

(84) P: ich glaube es ist vor allem Angst Erfolg zu haben (WIL.036.261)

(85) Amalie: Angst haben 2 (AMA.482)

```

01 P: wissen Sie
02 was ich immer versäumt und vermisst habe grad bei Träumen
03 dass ich da nie nach Herzenslust alles gesagt habe
04 was mir einfällt
05 weil ich immer Angst hatte und habe
06 glaub ich noch
07 Sie würden da entweder wachen oder böse werden
08 ich hab da immer
09 hinterher denk ich
10 und das könnte es noch sein und jenes
11 und da wird es derartig viel und bunt
12 ' s ganz , ganz dramatisch
13 sehr viel kommt da immer
14 und ich denke immer hier geht ' s noch recht karg zu

```



```

15      ' n roten Faden
16      zack fertig [...] Idylle

```

Unter Berücksichtigung des Ziels, Lexikoneinträge zu erstellen, die sich voneinander unterscheiden, resultieren aus der obigen Darstellung (mindestens) drei verschiedene Konstruktionen.⁴¹ Die möglichen adjektivischen und adverbialen Modifikationen (siehe Beispiel oben) sind in den folgenden Konstruktionen als Option enthalten und bilden keine eigenständige Konstruktion.

Tabelle 9.27: Konstruktion <Angst haben um>

Attribut	Wert	Bemerkung
canForm	<i>haben Angst um</i>	
crossRef	sich sorgen um	is-associated
entryFormation	phrase	
phraseType	set-phrase	
synFrame	550	AkkO + PräpO
synType	Funktionsverb	pseudotransitiv
semType	emotion	
definition	sich Sorgen machen um jemanden, Angst ausstehen	
natclause	jmd hat Angst um jmdn	
bauplan	550: 1 - - L2 subj Human 2 haben VAFIN L0 root n/a 3 Angst NN L2 obja n/a 4 um APPR L3 pp n/a 5 - - L4 pn Human	
pattern	[[Human]] haben {Angst} {um [[Human Anything]]}	
construction	[Angst haben um]	
subjField	Therapiegespräch	Diskurstyp
textType	unspezifisch	Diskurseinheit
register	unspezifisch	lokale Diskursmuster
topic	Besorgnis	lokales Pattern
Funktion	Kommunikative Funktion: seine Besorgnis ausdrücken, Teilnahme zeigen	
JAKOB-Code	SOR	SOR=Bangen (Wollen – Intendieren)

Es gibt sehr viele verwandte Konstruktionen zu den oben dargestellten drei Varianten von *Angst haben*. In erster Linie sind verschiedene Funktionsverbgefüge zu nennen, wie *Angst kriegen*, *Angst bekommen*, *Angst machen*, aber auch idiomatische Ausdrücke wie *Fracksausen haben*. Im VALBU ist die Konstruktion unter dem Verb *haben* generalisiert dargestellt. *Haben* (*Lesart 7*) ist hier so beschrieben: <jemand/etwas [Tier/Pflanze] empfindet

⁴¹Die Frage ist, ob *keine Angst haben* als Negation der Konstruktion *Angst haben* eine eigenständige Konstruktion darstellen soll. Die Erfassung in einer separaten Konstruktion ist vor allem für die JAKOB-Kodierung und -Analyse ein Thema, da für die positive und negative Variante evtl. andere JAKOB-Codes verwendet werden.

etwas [Gefühl]». Mit dieser Konstruktion können neben Angst ganz verschiedenartige Gefühle ausgedrückt werden.

9.3 Bemerkungen zu den Beispielkonstruktionen

Im Laufe der Arbeit an den Beispielen wurde ein neues Merkmal *Funktion* eingeführt, das aktuell noch nicht im Lexikon implementiert ist, das aber in die Tabellendarstellung der Konstruktionen aufgenommen wurde. Das Lexikon soll in Kürze mit diesem Merkmal erweitert werden (zusätzliches pragmatisches Merkmal). In wenigen Konstruktionsbeschreibungen wurde auch ein neues Merkmal *Modifikation* verwendet. Dieses Konzept wird jedoch nicht weiter verfolgt, weil in der OLIF-Struktur ein Merkmal *orthografische Variante* für verschiedene Lesarten vorgesehen ist.

Wie im Einführungsbeispiel zu den Lexikontabellen erwähnt wurde (9.1.2), erfordern die *JAKOB-Codes* eine vertiefte Kenntnis des Kodiersystems. Die JAKOB-Codes sind in den einzelnen Konstruktionsbeispielen enthalten, sie werden aber nicht detailliert hergeleitet. Die Kodierung ist nicht direkt abhängig von den übrigen Lexikonmerkmalen, d. h. es bestehen keine Regeln, wie die Codes aus den übrigen Merkmalen einer Konstruktion abgeleitet werden könnten.

Ausserdem sind die 93 Verbcodes in den meisten Fällen nicht ausreichend, um ähnliche Konstruktionen zu unterscheiden, wie aus verschiedenen Beispielen ersichtlich ist (vgl. die Konstruktionen <Sorge tragen zu> und <Sorge tragen für>, die beide mit SOR=angen kodiert werden). Die Kodierung ist nicht die beste Differenzierungsmöglichkeit für die Konstruktionen im vorliegenden Projekt, das Merkmal JAKOB-Code hat nicht diese Funktion. Auch vom methodischen Standpunkt her ist es sinnvoller, die Lexikon-Kategorien und die JAKOB-Kategorien im aktuellen Stadium des Projekts nicht zu vermengen, um die entstehenden Probleme besser zuordnen zu können.

Der aktuelle Status ist also der, dass Lexikoneinträge in der lexikographischen Arbeit mit linguistischen und extralinguistischen Eigenschaften versehen werden und dass die Zuweisung von JAKOB-Codes eine separate Expertenarbeit darstellt, die in einem zweiten Arbeitsgang durchgeführt wird. Die Qualität der computergestützten Textanalyse ist von der Qualität und vom Differenzierungspotenzial der einzelnen Konstruktionen abhängig, während die Qualität der Kodierung von der sachkundigen Zuordnung der

Codes zu den einzelnen Konstruktionen im Lexikon abhängig ist.

9.3.1 JAKOB-Lexikon online

Die aufgeführten Beispielkonstruktionen sollen Einblick geben in die lexicographische Arbeit, die gegenwärtig durchgeführt wird. Die beschriebenen Konstruktionen sind zum grössten Teil online zugänglich, neben vielen anderen. Das JAKOB-Lexikon umfasst aktuell (Dezember 2010) ca. 200 Einträge, die mit den oben vorgestellten Merkmalen versehen sind. Weitere 300 Einträge befinden sich in verschiedenen Stadien der Bearbeitung und dienen als prototypische Testfälle für die Anwendung und Weiterentwicklung der eingeführten Lexikonmerkmale.

URL: <http://www.jakob.uzh.ch/lexikon/>

10 Schluss: Fazit und Ausblick

Ziel der vorliegenden Arbeit war es, auf Grund verschiedener Einsichten aus Narrativik, Gesprächsforschung, Phraseologie, Konstruktionsgrammatik und Lexikographie ein bestehendes einfaches elektronisches Lexikon für die Kodierung von Erzähltexten auszubauen und mit Erweiterungen in Richtung eines *Konstruktionslexikons* zu versehen. Die Sichtung der genannten theoretischen Hintergründe hat wichtige Anregungen und Erkenntnisse zutage gefördert, um die Umsetzung zu einer praktisch einsetzbaren Lexikonressource in Angriff zu nehmen.

Was konnte umgesetzt werden? Was konnte implementiert, was nur angedacht werden? Wo ist weitere Arbeit nötig? Im Folgenden sollen die Erfahrungen kritisch beleuchtet und Vorschläge für die weitere Entwicklung des Lexikons gemacht werden. Konnten die in den Fragestellungen aufgeworfenen Fragen und Probleme beantwortet und gelöst werden? Was ist der Mehrwert des neuen des Lexikons für die Erzählanalyse JAKOB? Gibt es allenfalls weitere Verwendungsmöglichkeiten für das Lexikon?

Grundsätzlich gilt es bei allen Überlegungen zu beachten, dass auch das ausgefeilteste Lexikon nicht alle Bedeutungsvarianten erfassen kann; das mentale Lexikon der einzelnen Sprachbenutzer ist der Entwicklung immer voraus und kreiert ständig neue Bedeutungen.

Meaning does not reside in linguistic units, but is constructed in the minds of the language users. (Radden et al., 2007, S. 1)

10.1 Korpora

Einige kritische Anmerkungen müssen zu den verwendeten Korpora gemacht werden, und sie beantworten gleichzeitig die erste Forschungsfrage (S. 9): «Können *Konstruktionen* mit den vorgeschlagenen Methoden aus den Korpusbeispielen erschlossen und adäquat dargestellt werden, und welches sind die Anforderungen dafür?». Die drei verwendeten Korpora mit Therapiegesprächen (WIL, AMA, GUS) sind vom Umfang her sehr unterschiedlich (2.4 vs. 1.1 vs. 0.3 Millionen Tokens). In der Arbeit hat sich

herausgestellt, dass das Korpus WIL eine *Minimalgrösse* aufweist, mit der korpuslinguistische Fragen beantwortet werden können. Die Korpora AMA und GUS enthalten zwar für erzähl- und gesprächsanalytische Untersuchungen immer noch sehr viele Daten, aber die statistischen Assoziationsmassen sind unzuverlässig und von zufälligen Faktoren abhängig. Dies war der Hauptgrund dafür, die drei öffentlich zugänglichen Korpora AGD, CHTK und deWAC zusätzlich einzubeziehen. Die erste Forschungsfrage kann also dahingehend beantwortet werden, dass Konstruktionen aus Korpora differenziert erschlossen werden können, wenn die Korpusgrösse über 2 Millionen Tokens beträgt.

Das prominente Korpus WIL bildet Schweizer Dialekt ab und ist insofern ein Spezialfall (im Kontext der andern Korpora). Da wegen der erwähnten Korpusgrösse die Beispielkonstruktionen vor allem aus dem Korpus WIL stammen, ergibt sich dadurch eine Verzerrung, die im Kontext des Standarddeutschen störend ist, aber andererseits der Realität der mit der Erzählanalyse JAKOB untersuchten Erzählungen entspricht, die auf Transkripten aus dem Schweizer Sprachraum stammen. Insofern beruht eine Generalisierung auf den gefundenen Teilnehmer- und Sprecherkategorien und ist nicht von aussen herangetragen worden (Fischer, 2007).

Die drei zusätzlich hinzugezogenen Korpora haben dazu beigetragen, die Datengrundlage so zu vergrössern, dass sich das Übergewicht des Korpus WIL etwas relativierte. Das *Archiv für gesprochenes Deutsch* (AGD) ist insofern eine Rarität in der Korpuslandschaft, als dass es *gesprochene Sprache* beinhaltet und deshalb natürlich in die Recherche einbezogen werden musste. Die Arbeit mit diesem Korpus ist jedoch etwas umständlich und die Recherchemöglichkeiten linguistischer Art eher beschränkt.¹ Das *Schweizer Textkorpus* (CHTK) bietet ebenfalls etwas rudimentäre Recherchemöglichkeiten; der grösste Nachteil dieses Korpus besteht aber darin, dass die Daten auf geschriebenen Texten beruhen. Ausserdem sind viele der Texte sehr alt und in diesem Sinne nicht repräsentativ für das Lexikonvorhaben. Das deutsche Korpus, das in der *sketch engine* eingebaut ist, heisst *deWAC* und ist aus Internet-Texten verschiedenster Herkunft aufgebaut (Baroni & Kilgarriff, 2006). Die Recherche ist hier zugleich komfortabel und gut verständlich und ermöglicht richtiges *text mining*. Ein Vorteil ergibt sich weiter dadurch, dass die drei Therapiekorpora ebenfalls in der *sketch engine* erfasst sind und dadurch die Rechschritte analog ausgeführt werden können.

¹Dies könnte auch am Untersucher liegen, der nicht alle Möglichkeiten der Recherche über das COSMAS-Interface herausgefunden hat.

10.2 Das Konzept der Konstruktion

Die Forschungsfragen 2 bis 5 (S. 9) zielen auf die grundsätzliche Eignung des Konzepts *Konstruktion* für die Darstellung von lexikalischen Items und darauf, ob die einzelnen Konstruktionseigenschaften eine nützliche Differenzierung von Bedeutungsnuancen erlauben.

Dass das Konzept der *Konstruktion* geeignet ist für die Abbildung von Lexikoneinträgen wurde bereits mehrmals festgestellt. Diese Frage kann also mit *Ja* beantwortet werden. Konstruktionen umfassen von Einzelworten bis zu komplexen Wortverbindungen viele verschiedene Ausdruckskategorien. *Alle Merkmale* eines Lexikoneintrags charakterisieren gemeinsam die Konstruktion als Kombination von Form und Bedeutung, die mehr Informationen beinhaltet als ein formelhaftes Konstruktionsmerkmal.

Homonyme und polyseme Ausdrücke unterscheiden sich durch unterschiedliche Konstellationen von Form und Bedeutung, die von der Lexikographin entsprechend erfasst werden müssen, dann aber mindestens theoretisch eine merkmalsbasierte Unterscheidung zwischen mehreren ähnlichen Konstruktionen ermöglichen (Forschungsfrage 2). Die Forschungsfrage 3 (Kernbedeutungen im Lexikon) ist schwieriger empirisch zu beantworten. Es kann aber angenommen werden, dass das mentale Lexikon (individuelle) unterspezifizierte Kernbedeutungen enthält, die in der Interaktion mit zusätzlichen lokalen Bedeutungskomponenten versehen werden (Situations- und Kommunikationskontext). Dieses Modell kann als gedankliche Spekulation auf das JAKOB-Lexikon übertragen werden.

Die Forschungsfrage 4 (Wie viel Kontext müssen Lexikoneinträge beinhalten?) hängt mit der Frage 3 zusammen und ist ebenfalls nicht einfach zu beantworten. Unterspezifizierung bedeutet weniger Kontext und erlaubt grössere Freiheiten in der Anwendung von Lexikonwissen. Umfangreicher und präziser Kontext bedeutet aber eine Einschränkung des Bedeutungspotenzials. Für ein elektronisches Lexikon ist die letztgenannte Variante naheliegender.

Alle die genannten Aspekte der Konstruktion fließen in die folgende genauere Betrachtung einzelner wichtiger Lexikon-/Konstruktionsmerkmale ein.

10.3 Lexikonaufbau: Kritik und Vorschläge

Im Folgenden sollen die wichtigsten Erfahrungen und Probleme bei der Planung, Entwicklung und Implementierung der neuen Lexikonmerkmale

dargestellt werden, um daraus die Anforderungen für die Weiterentwicklung abzuleiten. Zu den wichtigsten neuen Lexikoneigenschaften gehören die *cross references*, die *Verbpatterns* und der *Bauplan*, und die *pragmatischen Attribute*.

10.3.1 Verknüpfungen (cross references)

Cross reference ist ein mächtiges Merkmal, das den Informationsgehalt eines Lexikoneintrags, besonders in einem elektronischen Lexikon, beträchtlich erweitern kann. Die OLIF-Struktur stellt dafür einen grossen Katalog von Verknüpfungsmöglichkeiten bereit (siehe Kapitel 9.1.2 und McCormick, 2005). Theoretisch könnte man also jede Konstruktion in einem grossen Netz von Knoten und Verzweigungen verorten, was konstruktionsgrammatischen Theorien entspricht (Goldberg, 1995) und was sicher auch einer Realität im mentalen Lexikon entspricht. Einige Verknüpfungen sind wichtig und treten in den Vordergrund, andere sind eher peripher und treten in den Hintergrund. Diese theoretisch sehr ansprechenden Möglichkeiten enden aber einstweilen vor der praktischen Umsetzung. Zwar könnten die Verknüpfungen wie in einem semantischen Netzwerk in die Datenbank implementiert werden, aber eine übersichtliche Darstellung ist im gegebenen Rahmen nicht möglich. Im aktuellen Layout der Lexikonansichten (Liste und einzelner Lexikoneintrag) werden zwar die Verknüpfungen angezeigt; mit mehreren Verknüpfungen wird aber der Datensatz rasch unübersichtlich. Dies könnte allenfalls so gelöst werden, dass man die Verknüpfungen gewichtet und nur die wichtigsten anzeigt.

Ein ebenso wichtiges Problem, das die Ausschöpfung des Potenzials der *cross references* behindert, ist der grosse lexikographische Aufwand, der für eine *systematische* Erfassung von Verknüpfungen geleistet werden muss. Wie bestimmt man die relevanten Verknüpfungen? Die Verknüpfungskonzepte müssten genauer definiert werden und auch die Ebenen, auf denen Verknüpfungen explizit angezeigt werden sollen (implizit gibt es in jedem Mehrwortausdruck Verbindungen zwischen den beteiligten Wortformen und Lexemen). Da die Kernbedeutung einer im Lexikon festgehaltenen Konstruktion sich je nach Kontext verändert, müssten gleichzeitig die Verknüpfungen ebenfalls anders gewichtet werden. Schliesslich gilt auch für die *cross references*, wie für Konstruktionen, dass Bedeutung auch dynamisch zustande kommen kann. So kann beispielsweise ein metonymisch verwendeter Ausdruck eine neue, ad hoc geschaffene Verknüpfung zu einem Konzept oder Lexikoneintrag herstellen.

Eine wichtige Aufgabe für die weitere Entwicklung des JAKOB-Lexikons

wird es sein, die Rolle der Verknüpfungen im Lexikon in zweierlei Hinsicht genauer festzulegen. Der erste Punkt betrifft die Anzahl der Verknüpfungen; schon bei einer sparsamen Zuordnung ergeben sich mehrere Verknüpfungen pro Eintrag. Wie viele davon sind notwendig und sinnvoll? Zum zweiten stellt sich die bis jetzt noch nicht beantwortete Frage, wie die Verknüpfungen bei der Suche im Lexikon berücksichtigt und bei der Textanalyse fruchtbar gemacht werden sollen.

10.3.2 Verbpatterns, Bauplan

Die wichtigste und, wie mir scheint, zukunftssträchtigste Neuerung im Lexikon ist die Kombination und das Zusammenspiel der Merkmale *natclause*, *bauplan* und *pattern*.

Verbpatterns sind Strukturmuster für das syntaktische und semantische Verhalten eines Verbs (Hanks, 2008a). Im Verbpattern nach dem Vorbild der CPA (Corpus Pattern Analysis) sind die semantischen Eigenschaften aller Ergänzungen und Angaben einer Verbalphrase abgebildet, und zwar nach Massgabe ihres empirisch erschlossenen Vorkommens in Textkorpora. Gemäss Hanks gibt es (für die englische Sprache) für jedes voll spezifizierte Verbpattern eine bestimmte Bedeutung. Die Stärke der Verbpatterns für das vorliegende Projekt liegt darin, dass alle im Zusammenspiel mit einem bestimmten Verb in einer bestimmten Kommunikationssituation relevanten Satzkomponenten mit semantischen Typen (und Rollen) versehen sind und so eine möglichst eindeutige Konstellation bilden.

Das Merkmal *natclause* dient zur Generierung eines maschinenlesbaren Merkmals *bauplan* mit Hilfe eines Parsers (siehe Kapitel 9.1.3). Das Merkmal *bauplan* bildet alle syntaktischen und semantischen Merkmale und Restriktionen einer Konstruktion in maschinenlesbarer Form ab und bildet damit die für die Analyseapplikation verwertbare Entsprechung des Verbpatterns. Dies ermöglicht es, die potentielle Satzstruktur eines Lexikoneintrags zu verwenden, um bei der maschinellen Textanalyse diese Struktur mit der Struktur eines Satzes (Segments) aus dem zu analysierenden Text zu vergleichen und auf Übereinstimmung zu prüfen.

Die drei genannten Merkmale bilden zusammen das syntaktisch-semantische Herzstück der Konstruktion, in einer intuitiv und leicht verständlichen Form für die Lexikonbenutzer in Form des *patterns*, und einer Variante für die computergestützte Textanalyse (*bauplan*). Verbale Konstruktionen können auf diese Art mit einer optimal differenzierenden Struktur dargestellt werden.

10.3.3 Pragmatische Attribute

Die neu eingeführten pragmatischen Merkmale *subject field*, *text type*, *register* und *topic* sind hingegen eher die Sorgenkinder der neuen Lexikonversion; sie bilden Eigenschaften der Dikurssituation ab, die nur teilweise im Lexikon abgebildet werden können, weil Lexikoneinträge bis zu einem gewissen Grade dekontextualisiert sind.

Das Merkmal *subject field* (auch Domain oder Genre) könnte eigentlich im vorliegenden Kontext immer auf *Therapiegespräch* gesetzt werden, was aber für viele Lexikoneinträge eine etwas gesuchte Einschränkung wäre. Dasselbe gilt für das Merkmal *text type*, das lokale linguistische Merkmale der Gesprächssituation bezeichnet, wie *narrativer* oder *argumentativer* Gesprächsstil, *Beschreibung* etc. Auch hier sind die Fälle selten, bei denen eine feste Zuordnung im Lexikon vorgenommen werden kann. Am besten für eine Spezifizierung im Lexikon eignet sich das Merkmal *register*, mit dem Konstruktionen als *umgangssprachlich*, *gehoben* oder *salopp* bezeichnet werden können.

Das Merkmal *topic* ist in der aktuellen Implementierung noch nicht präzise definiert. Es ermöglicht in gewissen Fällen, Konstruktionen in einem lokalen *inhaltlichen Kontext* zu verorten, so wenn sich die Gesprächseinheit auf praktische Verrichtungen und Situationen aus dem Alltag bezieht, wie Ferien, Wandern, Autofahren etc. Die möglichen Kategorien für dieses Merkmal werden nach Bedarf ergänzt und geändert. Eine Ergänzungsmöglichkeit, die zu prüfen ist, könnte darin bestehen, das Merkmal *topic* mit den *Frames* (sensu FrameNet) in Verbindung zu bringen.

Die Frage nach der Zuordnung der pragmatischen Kategorien zu den Konstruktionen hat sich als schwierig erwiesen. Eigentlich müssten diese als Eigenschaften der *Gesprächssituation* bei der lexikalischen Analyse als paralleles *Situationswissen* neben dem *Lexikonwissen* mitgeführt werden und in die Bedeutungszuweisung einfließen. So können z. B. saloppe Redensarten in gewissen Gesprächskontexten als unangebracht und aggressiv empfunden werden, während sie in einem andern Gesprächskontext als Norm gelten.

10.3.4 Diskurs über den Satz hinaus

Wilma verwendet einmal das Sprichwort: «Wenn die Katze aus dem Haus ist, tanzen die Mäuse.» Damit stellt sich die Frage, wie grössere Diskursseinheiten behandelt werden sollen. Das zitierte Sprichwort kann natürlich als Konstruktion angesehen werden, und es spricht nichts dagegen, Konstruktionen als Gebilde zu definieren, die über die Satzgrenze hinausgehen

können. Da das Lexikon aber – auf Grund der Verbpatterns und Satzmuster, aber besonders weil die Analyseapplikation auf Segmentebene operiert – satzorientiert ist, gibt es aktuell keine konzeptuell passende und nahe liegende Möglichkeit, Konstruktionen grösser als Sätze ins Lexikon aufzunehmen.

10.4 Dynamische Bedeutungskonstitution

Ein grundsätzlicher Aspekt von Konstruktionen, der sich in der ganzen Arbeit immer wieder gezeigt hat, ist die Frage nach der *Bedeutungskonstitution* (oder *meaning construction*). Neben der im Lexikon abgebildeten statischen Liste von Konstruktionen, gibt es immer wieder neu gebildete Konstruktionen. Wie werden Konstruktionsbedeutungen prozessiert, die im Sprachgebrauch kreativ neu erfunden wurden? Wie funktioniert hier das Meinen und Verstehen? Können vorhandene Konstruktionen im Lexikon und ihre Bedeutungspotenziale etwas zur Erschliessung «neuer» Bedeutung beitragen (vgl. Fragestellung 6)? Dabei geht es nicht nur um die Aspekte der Bedeutungskonstitution, die praktisch mit dem Schreiben eines Lexikons zusammenhängen, sondern auch um die theoretischen Modelle, die die Prozesse der Bedeutungskonstitution und des *meaning construction* erklären.

Linell zitiert dazu eine *interaktionale Sicht* auf Konstruktionen (nach Anward, forthcoming) und formuliert die folgenden fünf Punkte (Linell, 2009, S. 311):

- Konstruktionen sind *Äusserungstypen*, Abstraktionen von konkreten Äusserungen oder Ausschnitte aus Netzwerken von verknüpften Äusserungen.
- Konstruktionen sind oft *lexikalisiert*, haben aber trotzdem *variable* Teile.
- Konstruktionen sind *offen und dynamisch*, sowohl paradigmatisch als auch syntagmatisch; sie können in verschiedene Richtungen angepasst werden.
- Konstruktionen und ihre Instanziierungen sind sowohl *kontext-abhängig* als auch *kontext-einführend*.
- Konstruktionen sind gleichzeitig *abstrakte Formeln*, aus dem Kontext entfernt, d. h. dekontextualisiert, und tragen trotzdem einen Teil ihrer

«alten» *Kontexte* mit sich, wenn sie in «neue» *Kontexte* eingebettet werden.

10.4.1 Meaning Construction

Meaning construction wurde bereits im Kapitel (3.3) vorgestellt und wird bei Radden et al. (2007, S. 3) als eine ständig ablaufende mentale Aktivität bezeichnet, bei der die Gesprächsteilnehmer in jedem kommunikativen Akt Bedeutung(en) auf der Basis von unterspezifizierten linguistischen Einheiten (z. B.. aus Grammatik und Lexikon) herstellen. *Meaning construction* betrifft nicht nur «neue» Bedeutung; d. h. nicht nur neuartige Bedeutung wird konstruiert, sondern jede Bedeutung in jeder Situation wird bis zu einem gewissen Grad neu ausgehandelt, sowohl auf der produktiven als auch auf der rezeptiven Ebene. *Meaning construction* ist eine *iterative Prozedur*, die gelingen oder scheitern kann. Wichtig für die Erzähl- und Gesprächsanalytikerin wäre es, die Mechanismen des *meaning construction* zu kennen und bei der Analyse zu berücksichtigen.

Die Prozesse des *meaning construction* werden gemäss Radden et al. (2007, S. 5) vor allem über das von ihnen als äusserst wichtig erachtete Phänomen der *Unterspezifizierung* ermöglicht. Es gibt drei Typen von *Unterspezifizierung* (siehe Kapitel 3.3.1).

- *Unterspezifizierung* durch Implicitness
- *Unterspezifizierung* durch Nicht-Determiniertheit
- *Unterspezifizierung* durch Inkompatibilität

Die wichtige Rolle, die *Unterspezifizierung* und *Vagheit* bei der Bedeutungskonstitution spielen, wird auch von andern Autoren betont (Deignan, 2005; Feilke, 1996; Linell, 2009), die zeigen, wie wichtig systematische Vagheit bei der Produktion und Rezeption von idiomatischen, metaphorischen oder metonymischen Ausdrücken ist. Metonymische und metaphorische Aussagen evozieren vage Vorstellungen und Stereotypen, die in der Rezeption und Reproduktion zu neuen kreativen Bildern und Konstruktionen führen. Kommunikatives Ziel des Metonymie- und Metapherngebrauchs ist nicht die kognitive Präzision, sondern Vagheit. Die Sprache als Kommunikationsinstrument erfordert Vagheit; die semantische Plastizität von Sprache ist eine notwendige Bedingung der Kommunikation (Feilke, 1996). Das Phänomen der Vagheit wurde z. B.. anhand der Konstruktion <an den Karren

fahren› oder anhand der ‹Zeug›-Konstruktionen gezeigt (Kapitel 9.2.2 und 9.2.4).

Ein wichtiges Ziel für die Weiterentwicklung des Lexikons wird sein, diese Grundsätze im Auge zu behalten und immer wieder zu überlegen, wie kreativer Sprachgebrauch, Unterspezifizierung und Vagheit mit den Lexikonkonzepten in Verbindung gebracht werden können, als Ergänzung oder möglicherweise als Alternative zu den lexikonspezifischen Bemühungen um Kategorisierung.

10.4.2 Norms and Exploitations

Interessante Aspekte der Bedeutungsherstellung beschreibt auch die im Kapitel zur *Corpus Pattern Analysis* vorgestellte Theorie der *norms and exploitations* (Hanks, 2004, in press). Sprachkompetenz erfordert einerseits die Fähigkeit, Ausdrücke regelhaft im wörtlichen und idiomatischen Sinn anzuwenden (*norms*), als auch die Kompetenz, kreativ neue und von diesen Regeln abweichende Verwendungsarten dieser Ausdrücke zu erkennen oder selbst zu erfinden (*exploitations*). So entspricht die Konstruktion ‹Dampf ablassen› aus dem Korpus WIL der *Norm* gemäss Sprachgebrauch in diversen Korpora, während es sich bei der Konstruktion ‹Dampf aufbauen› um eine *exploitation* handelt (siehe Kapitel 9.2.5).

Ein hypothetischer Verarbeitungsablauf am Beispiel von ‹Dampf aufbauen› könnte so aussehen:

1. Vorausgesetzt: Kenntnis des Idioms ‹Dampf ablassen› (im Lexikon).
2. Implikatur: Aggression senken, seine Wut reduzieren.
3. Der Gegensatz zu *ablassen* ist *aufbauen* (Antonym). Dies geht ebenfalls aus dem Lexikon hervor (Gegensatz von *verkleinern* – *vergrößern*).
4. D. h. die gewünschte Implikatur ist: Antonym zu Dampf ablassen
5. Draus folgt die Kreation der Konstruktion ‹Dampf aufbauen› (*exploitation*).

10.4.3 Das Paradoxon des Formelhaften

Gesprächstranskripte (reale Sprachdaten) zeigen, dass *Formelhaftes* einen wesentlichen Teil unserer Sprachverwendung ausmacht. Formelhaftes zeigt

sich in verfestigten Ausdrücken und Strukturen (Idiome, Metaphern, Metonymien, Sprichwörter), aber auch in konventionalisierten Wortverbindungen.

Im Gegensatz dazu und als Paradoxon stellt man aber anhand der Transkripte fest, dass das Formelhafte auch Gegenstand von häufigen *Variationen* und Veränderungen ist. Einerseits werden formelhafte Ausdrücke so variiert, dass sie beispielsweise gewisse Bedeutungsaspekte in den Vordergrund stellen und andere zurückdrängen, andererseits gibt es neue Wort- und Ausdruckskombinationen, die im Gespräch bewusst neu eingeführt oder irrtümlich, z. B.. durch die «falsche» Anwendung einer Metapher, eines Idioms oder einer festen Wortverbindung, verwendet werden, im Kontext trotzdem verständlich sind und auch den weiteren Kontext wieder mitbestimmen.

Nun könnte man einwenden, dass der Begriff «Formelhaftes» nicht am Platz ist, wenn doch die Formeln regelmässig umgebaut werden. Steyer und Brunner (2009) sagen zur dynamischen Aktualisierung von *Wortbindungsmustern*:

Die Ebene «Wortbindungsmuster» [...] bildet [sie] kognitive Mechanismen des Sprechens und Verstehens ab, und zwar das Phänomen der dynamischen Aktualisierung von konkreten Gebrauchsbedeutungen einer Wortverbindung und dem zu Grunde liegenden Muster in einer Äusserungssituation. Die Fähigkeit von Sprechern, konkrete Äusserungen vor der Folie eines Musters höherer Ordnung zu verstehen, wird in der Phraseologie vor allem als Erklärung für das Verstehen von sprachspielerischen Modifikationen herangezogen. Sie scheint aber sehr viel weitreichender für das Funktionieren von Kommunikation zu sein und eben nicht nur im engen Sinne für den Gebrauch von Phraseologismen zu gelten. [...] Die Sprecher können diese Vagheitsfunktion sowohl einsetzen als auch decodieren, weil sie viele Verwendungsfälle ähnlicher Art bereits als verfestigte Muster im Kopf haben. (S. 35)

Parallel zur Weiterentwicklung des Lexikons sollten weitere Studien zur formelhaften Sprache und ihrem Verhältnis zum Lexikon durchgeführt werden (vgl. Wray, 2005, 2008 und Kapitel 4.7).

10.5 Erzählanalyse JAKOB – Computeranwendung

Die letzte Fragestellung (7) betrifft die Frage, ob die vorgenommenen Lexikonergänzungen die *automatische Kodierung* in der Erzählanalyseapplikation verbessern. Nach den Neuerungen und Ergänzungen im Lexikon besteht die Folgeaufgabe darin, die *Analyseapplikation* den neuen Gegebenheiten anzupassen. Ein erster Schritt dazu ist bereits in Angriff genommen worden: Da durch die neuen Lexikonmerkmale auch neue Anforderungen an die Kodierprozedur entstehen, muss diese vollständig neu implementiert werden. Die *Neuimplementierung* wird parallel zur bestehenden Applikation aufgebaut, so dass kein Betriebsunterbruch entsteht.

Eine wesentliche und sinnvolle Neuerung wird darin bestehen, dass der *Parser*², der jetzt verwendet wird, um die syntaktischen Strukturen für das Merkmal *bauplan* zu generieren, auch für das Parsing der zu analysierenden Erzähltexte verwendet wird. Da die Erzähltexte für die Analyse manuell segmentiert werden, ergibt sich für den Parser eine ähnliche Aufgabe wie beim Parsing der Mustersätze aus dem Merkmal *natclause*, ausser dass diese wohlgeformter sind. Dadurch können im günstigen Fall, wenn der Parser syntaktische Strukturen oder Teilstrukturen erkennt, syntaktische Rollen und Verbmuster abgeleitet werden. Dies steht im Gegensatz zum bis anhin verfolgten *bag-of-words*-Ansatz (BOW) (Erk & Padó, 2010), der für die JAKOB-Analyse darin besteht, dass alle Wörter eines Segments als potentiell gleich bedeutungstragend eingeschätzt und unstrukturiert abgearbeitet werden.

Die *Segmentierung* der Gesprächstexte ist allerdings nicht unumstritten. Für die computergestützte Textanalyse werden die Texte oder Gesprächstranskripte üblicherweise nach den strukturalistischen Kriterien wie Satz, Redeteil, Subjekt, Prädikat etc. segmentiert. Argelander (1991) zeigt ebenfalls mit seinen Begriffen der *Aussagen* oder *Sätze*, dass die Segmentierung eine plausible Methode der Textstrukturierung ist. Die Frage ist, ob diese Art Segmentierung der Struktur der gesprochenen Sprache gerecht wird. Auer (2010) bezweifelt dies und argumentiert, dass der prozesshafte (emergente) Charakter der Sprachproduktion durch die Segmentierung durchbrochen wird. Er schlägt vor, das Segmentieren durch das Zäsurieren zu ersetzen, das auf Abschlusspunkte ausserhalb der Satzstruktur zielt. Diese Alternativen können aber für NLP-Anwendungen nicht verwendet werden.

²Pro3GresDe – a Dependency Parser for German.

URL: <https://files.ifi.uzh.ch/cl/gschneid/parser/>

Für die Computerbearbeitung von elektronischen Texten ist die Segmentierung im Sinne einer Satzstruktur die einzige gangbare Möglichkeit.

Es bleibt die Frage, wie die computergestützte Erzählanalyse mit dem kreativen Sprachgebrauch umgehen soll. Dafür wird es wohl in absehbarer Zeit keine Lösung geben; dieses Feld wird den Erzähl- und Gesprächsanalytikern noch für längere Zeit überlassen bleiben.

11 Literaturverzeichnis

- Agirre, E. & Edmonds, P. (2006a). Introduction. In E. Agirre & P. Edmonds (Hrsg.), *Word sense disambiguation* (S. 1–28). Dordrecht: Springer.
- Agirre, E. & Edmonds, P. (Hrsg.). (2006b). *Word sense disambiguation: Algorithms and applications*. Dordrecht: Springer.
- Aijmer, K. & Stenström, A.-B. (2005). Approaches to spoken interaction. *Journal of Pragmatics*, 37 (11), 1743–1751.
- Albani, C., Pokorny, D., Blaser, G. & Kächele, H. (2008). *Beziehungsmuster und Beziehungskonflikte: Theorie, Klinik und Forschung*. Göttingen: Vandenhoeck & Ruprecht.
- Angus, L. E. & McLeod, J. (Hrsg.). (2004). *The handbook of narrative and psychotherapy. Practice, theory, and research*. Thousand Oaks: Sage.
- Antaki, C. (2007). Mental-health practitioners' use of idiomatic expressions in summarising clients' accounts: Special Issue: Diversity and Continuity in Conversation Analysis. *Journal of Pragmatics*, 39 (3), 527–541.
- Antaki, C., Barnes, R. & Leudar, I. (2005). Diagnostic formulations in psychotherapy. *Discourse Studies*, 7 (6), 627–647.
- Arboleda, L., Boothe, B., Grimm, G., Hermann, M.-L., Luder, M., Neukom, M. et al. (2010). *Kurzanleitung zur Erzählanalyse JAKOB: Version 10/10*. Zürich.
- Argelander, H. (1991). *Der Text und seine Verknüpfungen: Studien zur psychoanalytischen Methode*. Berlin: Springer.
- Auer, P. (2003). Realistische Sprachwissenschaft. In A. Linke, H. Ortner & P. R. Portmann-Tselikas (Hrsg.), *Sprache und mehr* (Bd. 245, S. 177–188). Tübingen: Niemeyer.
- Auer, P. (2006). Construction Grammar meets Conversation Analysis: Einige Überlegungen am Beispiel von "so"-Konstruktionen. In S. Günthner & W. Imo (Hrsg.), *Konstruktionen in der Interaktion* (Bd. 20, S. 291–314). Berlin: de Gruyter.
- Auer, P. (2007). Syntax als Prozess. In H. Hausendorf (Hrsg.), *Gespräch als Prozess* (Bd. 37, S. 95–124). Tübingen: Gunter Narr.
- Auer, P. (2010). Zum Segmentierungsproblem in der Gesprochenen Sprache. *InLiSt - Interaction and Linguistic Structures* (49), 1–19. Verfügbar unter <http://www.inlist.uni-bayreuth.de/issues/49/InLiSt49.pdf>
- Auer, P. & Günthner, S. (2003). Die Entstehung von Diskursmarkern im Deutschen - ein Fall von Grammatikalisierung? *InLiSt - Interac-*

- tion and Linguistic Structures, 38 (Dezember 2003). Verfügbar unter <http://www.uni-potsdam.de/u/inlist/issues/38/>
- Baker, C. F., Fillmore, C. J. & Cronin, B. (2003). The Structure of the Framenet Database. *International Journal of Lexicography*, 16 (3), 281–296.
- Ballmer, T. T. & Brennenstuhl, W. (1986). *Deutsche Verben: Eine sprachanalytische Untersuchung des deutschen Verbwortschatzes* (Bd. 19). Tübingen: Narr.
- Bamberg, M. (2008). Narrative Analysis and Identity Research: A Case for 'Small Stories'. *Manuscript submitted for publication*. Verfügbar unter <http://www.clarku.edu/~mbamberg/publications.html>
- Bamberg, M. & Georgakopoulou, A. (2008). Small stories as a new perspective in narrative and identity analysis. *Text & Talk - An Interdisciplinary Journal of Language, Discourse & Communication Studies*, 28 (3), 377–396.
- Baroni, M. & Kilgarrieff, A. (2006). Large linguistically-processed Web corpora for multiple languages. *Proceedings EACL 2006*, 1–4.
- Belica, C. (2008). Semantische Nähe als Ähnlichkeit von Kookkurrenzprofilen. *Korpusinstrumente in Lehre und Forschung*, 1–16.
- Bercelli, F., Rossano, F. & Viaro, M. (2008). Different place, different action: Clients' personal narratives in psychotherapy. *Text & Talk - An Interdisciplinary Journal of Language, Discourse & Communication Studies*, 28 (3), 283–305.
- Bergen, B. K. & Chang, N. (2005). Embodied Construction Grammar in simulation-based language understanding. In J.-O. Östman & M. Fried (Hrsg.), *Construction grammars* (Bd. 3, S. 147–190). Amsterdam: Benjamins.
- Bergmann, J. & Luckmann, T. (1995). Drama and Narration. In U. M. Quasthoff (Hrsg.), *Aspects of oral communication* (Bd. 21, S. 289–304). Berlin etc.: de Gruyter.
- Bergs, A. & Diewald, G. (Hrsg.). (2008a). *Constructions and language change* (Bd. 194). Berlin: Mouton de Gruyter.
- Bergs, A. & Diewald, G. (2008b). Introduction: Constructions and Language Change. In A. Bergs & G. Diewald (Hrsg.), *Constructions and language change* (Bd. 194, S. 1–22). Berlin: Mouton de Gruyter.
- Berth, H., Balck, F. & Brähler, E. (2005). Sprachinhaltsanalyse in der kulturvergleichenden psychosomatischen Forschung: Die Gottschalk-Gleser-Angstskalen. *Psychotherapie Psychosomatik Medizinische Psychologie*, 55 (12), 493–501.
- Biber, D. & Finegan, E. (1994a). Introduction: Situating register in sociolinguistics. In D. Biber & E. Finegan (Hrsg.), *Sociolinguistic perspectives on register* (S. 3–12). New York: Oxford University Press.
- Biber, D. & Finegan, E. (Hrsg.). (1994b). *Sociolinguistic perspectives on register*. New York: Oxford University Press.
- Boas, H. C. (2005). From theory to practice: Frame semantics and the design of

- FrameNet. In S. Langer & D. Schnorbusch (Hrsg.), *Semantik im Lexikon* (S. 129–160). Tübingen: Narr.
- Boothe, B. (1989). *Zur psychoanalytischen Konfliktdiagnostik: Entwicklung eines hermeneutischen Verfahrens zur diagnostischen Auswertung von Erstinterview- und Therapieprotokollen*. Bern: Peter Lang.
- Boothe, B. (2002). *Kodiermanual zur Erzählanalyse JAKOB. Version 10/02* (Nr. 51a). Universität Zürich.
- Boothe, B. (2004). *Der Patient als Erzähler in der Psychotherapie* (2. Aufl.). Gießen: Psychosozial-Verlag.
- Boothe, B. (2009). Erzählen im medizinischen und psychotherapeutischen Diskurs. In C. Klein & M. Martinez (Hrsg.), *Wirklichkeitserzählungen* (S. 51–80). Stuttgart: Metzler.
- Boothe, B. (2011). *Das Narrativ: Biografisches Erzählen im psychotherapeutischen Prozess*. Stuttgart: Schattauer.
- Boothe, B., Grimm, G., Hermann, M.-L. & Luder, M. (2010). JAKOB Narrative Analysis: The psychodynamic conflict as a narrative model. *Psychotherapy Research*, 20 (5), 511–525.
- Boothe, B., Grimmer, B., Luder, M., Luif, V., Neukom, M. & Spiegel, U. (2002). *Manual der Erzählanalyse JAKOB: Version 10/02* (Nr. 51). Zürich.
- Braasch, A. & Pedersen, B. S. (2010). Encoding Attitude and Connotation in wordnets. In A. Dykstra & T. Schoonheim (Hrsg.), *Proceedings of the XIV Euralex International Congress* (S. 282–292).
- Brändle, J. (2008). *Träume erzählen in der Psychotherapie: Eine erzählanalytische Untersuchung der Träume von Frau W.* Unveröffentlichte Lizentiatsarbeit, Universität Zürich, Psychologisches Institut, Klinische Psychologie, Psychotherapie und Psychoanalyse.
- Brändle, J. (2009). Träume erzählen in der Psychotherapie - Eine erzählanalytische Untersuchung der Träume von Frau W. In G. Grimm, N. Kapfhammer, H. Mathys, S. Michel & B. Boothe (Hrsg.), *Erzählen, Träumen, Erinnern. Erträge Klinischer Erzählforschung* (S. 180–187). Lengerich: Pabst.
- Brinker, K. & Sager, S. F. (2006). *Linguistische Gesprächsanalyse: Eine Einführung* (4. Aufl.). Berlin: Erich Schmidt.
- Bruner, J. S. (1986). *Actual minds, possible words*. Cambridge, MA: Harvard University Press.
- Bruner, J. S. (1990). *Acts of Meaning*. Cambridge, MA: Harvard University Press.
- Bruner, J. S. (2004). The narrative creation of self. In L. E. Angus & J. McLeod (Hrsg.), *The handbook of narrative and psychotherapy. Practice, theory, and research* (S. 3–14). Thousand Oaks: Sage.
- Buchholz, M. B. (2009). Spiegelneuronengedächtnispolitik: Ein kleiner Literaturrundflug. *DGPT Psycho-News-Letter* (75), 1–20.
- Burchadt, A., Erk, K., Frank, A., Kowalski, A., Padó, S. & Pinkal, M. (2009). FrameNet for the semantic analysis of German: Annotation, representa-

- tion and automation. In H. C. Boas (Hrsg.), *Multilingual FrameNets in Computational Lexicography* (S. 209-244). de Gruyter.
- Burger, H. (2004). Phraseologie - kräuter und rüben? traditionen und perspektiven der forschung. In K. Steyer (Hrsg.), *Wortverbindungen - mehr oder weniger fest* (Bd. 2003, S. 19-40). Berlin: de Gruyter.
- Burger, H. (2007). *Phraseologie: Eine Einführung am Beispiel des Deutschen* (3., neu bearb. Aufl. Aufl.). Berlin: E. Schmidt.
- Burger, H., Dobrovol'skij, D., Kühn, P. & Norrick, N. R. (Hrsg.). (2007). *Phraseologie: Ein internationales handbuch der zeitgenössischen forschung*. Berlin: W. de Gruyter.
- Bybee, J. (2006). From Usage to Grammar: The Mind's Response to Repetition. *Language*, 82 (4), 711-733.
- Carstensen, K.-U., Ebert, C., Endriss, C., Jekat, S., Klabunde, R. & Langer, H. (Hrsg.). (2004). *Computerlinguistik und Sprachtechnologie: Eine Einführung* (2. Aufl.). München: Elsevier.
- Church, K. W. & Hanks, P. (1990). Word Association Norms, Mutual Information, and Lexicography. *Computational Linguistics*, 16 (1), 22-29.
- Cohn, M. A., Mehl, M. R. & Pennebaker, J. W. (2004). Linguistic Markers of Psychological Change Surrounding September 11, 2001. *Psychological Science*, 15 (10), 687-693.
- Copestake, A., Lambeau, F., Villavicencio, A., Bond, F., Baldwin, T., Sag, I. et al. (2002). Multiword expressions: Linguistic precision and reusability. *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC 2002)*, Las Palmas, 1941-1947.
- Couper-Kuhlen, E. & Selting, M. (2001). Introducing Interactional Linguistics. In M. Selting & E. Couper-Kuhlen (Hrsg.), *Studies in interactional linguistics* (Bd. 10, S. 1-22). Amsterdam: John Benjamins.
- Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Croft, W. & Cruse, D. A. (2004). *Cognitive linguistics* (3. Aufl.). Cambridge: Cambridge University Press and Univ. Press.
- Cruse, D. A. (2004). *Meaning in language: An introduction to semantics and pragmatics* (2. Aufl.). Oxford: Oxford University Press.
- Dausendschön-Gay, U. & Gülich, E. (2007). Vorgeformtheit als Ressource im konversationellen Formulierungs- und Verständigungsprozess. In H. Hausendorf (Hrsg.), *Gespräch als Prozess* (Bd. 37, S. 181-219). Tübingen: Gunter Narr.
- Deignan, A. (2005). *Metaphor and corpus linguistics* (Bd. vol. 6). Amsterdam: Benjamins.
- Deppermann, A. (2005). Conversational interpretation of lexical items and conversational contrasting. In A. Hakulinen & M. Selting (Hrsg.), *Syntax and lexis in conversation* (Bd. Vol. 17, S. 289-317). Amsterdam: John Benjamins.
- Deppermann, A. (2006a). Construction Grammar - Eine Grammatik für die

- Interaktion? In A. Deppermann, R. Fiehler & T. Spranz-Fogasy (Hrsg.), *Grammatik und Interaktion* (S. 43–65). Radolfzell: Verlag für Gesprächsforschung.
- Deppermann, A. (2006b). Deontische Infinitivkonstruktionen: Syntax, Semantik, Pragmatik und interaktionale Verwendung 239. In S. Günthner & W. Imo (Hrsg.), *Konstruktionen in der Interaktion* (Bd. 20, S. 239–262). Berlin: de Gruyter.
- Deppermann, A. (2006c). *Gespräche analysieren: Eine Einführung* (4. Aufl., Bd. 3). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Deppermann, A. (2006d). Konstitution von Wortbedeutung im Gespräch: Eine Studie am Beispiel des jugendsprachlichen Bewertungsadjektivs *assi*. In A. Deppermann & T. Spranz-Fogasy (Hrsg.), *Be-deuten* (S. 158–184). Tübingen: Stauffenburg.
- Deppermann, A. (2006e). Von der Kognition zur verbalen Interaktion: Bedeutungskonstitution im Kontext aus der Sicht der Kognitionswissenschaften und der Gesprächsforschung. In A. Deppermann & T. Spranz-Fogasy (Hrsg.), *Be-deuten* (S. 11–33). Tübingen: Stauffenburg.
- Deppermann, A. (2007). *Grammatik und Semantik aus gesprächsanalytischer Sicht* (Bd. 14). Berlin: de Gruyter.
- Deppermann, A. & Elstermann, M. (2008). Lexikalische Bedeutung oder Konstruktionsbedeutungen? Eine Untersuchung am Beispiel von Konstruktionen mit "verstehen". In A. Stefanowitsch & K. Fischer (Hrsg.), *Konstruktionsgrammatik II* (Bd. 47, S. 103–134). Tübingen: Stauffenburg.
- Deppermann, A., Fiehler, R. & Spranz-Fogasy, T. (Hrsg.). (2006). *Grammatik und Interaktion: Untersuchungen zum Zusammenhang von grammatischen Strukturen und Gesprächsprozessen*. Radolfzell: Verlag für Gesprächsforschung.
- Deppermann, A. & Spranz-Fogasy, T. (2001). Aspekte und Merkmale der Gesprächssituation. In K. Brinker, G. Antos, W. Heinemann & S. F. Sager (Hrsg.), *Text- und Gesprächslinguistik* (Bd. 2. Halbband, S. 1148–1160). Berlin: de Gruyter.
- Dijk, T. A. van. (2008). *Discourse and context: A sociocognitive approach*. Cambridge: Cambridge University Press.
- Dobrovol'skij, D. (2004). Idiome aus kognitiver Sicht. In K. Steyer (Hrsg.), *Wortverbindungen - mehr oder weniger fest* (Bd. 2003, S. 117–143). Berlin: de Gruyter.
- Dornseiff, F. (2004). *Der deutsche Wortschatz nach Sachgruppen* (8. Aufl.). Berlin: de Gruyter.
- Du Bois, J. W. (2003). Discourse and Grammar. In M. Tomasello (Hrsg.), *The new psychology of language* (Bd. 2, S. 47–88). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Du Bois, J. W., Kumpf, L. E. & Ashby, W. J. (Hrsg.). (2003). *Preferred argument structure: Grammar as architecture for function* (Bd. 14). Amsterdam:

- Benjamins.
- Duffner, R. & Näf, A. (2006). Digitale Textdatenbanken im Vergleich. *Linguistik Online*, 28 (3), 7-22.
- Engelberg, S. & Lemnitzer, L. (2009). *Lexikographie und Wörterbuchbenutzung* (4., überarb. und erw. Aufl. Aufl., Bd. 14). Tübingen: Stauffenburg-Verl.
- Erk, K. & Padó, S. (2010). Exemplar-Based Models for Word Meaning In Context. *Proceedings of the ACL 2010 Conference, Uppsala*.
- Evert, S. (2005). *The Statistics of Word Cooccurrences: Word Pairs and Collocations: Dissertation*. Stuttgart.
- Fauconnier, G. (1994). *Mental spaces: Aspects of meaning construction in natural language*. Cambridge: Cambridge University Press.
- Feilke, H. (1996). *Sprache als soziale gestalt: Ausdruck, prägung und die ordnung der sprachlichen typik*. Frankfurt a. M.: Suhrkamp.
- Feilke, H. (2004). Kontext - zeichen - kompetenz: Wortverbindungen unter sprachtheoretischem aspekt. In K. Steyer (Hrsg.), *Wortverbindungen - mehr oder weniger fest* (Bd. 2003, S. 41-64). Berlin: de Gruyter.
- Feilke, H. (2007). Syntaktische Aspekte der Phraseologie III: Construction Grammar und verwandte Ansätze. In H. Burger, D. Dobrovolskij, P. Kühn & N. R. Norrick (Hrsg.), *Phraseologie* (S. 63-76). Berlin: W. de Gruyter.
- Fellbaum, C. (1998a). A Semantic Network of English: The Mother of All Word-Nets. *Computers and the Humanities*, 32 (2-3), 209-220.
- Fellbaum, C. (Hrsg.). (1998b). *WordNet: An electronic lexical database*. Cambridge, Mass.: MIT Press.
- Fellbaum, C. (2007a). Argument selection and alternations in vp idioms. In C. Fellbaum (Hrsg.), *Idioms and collocations* (S. 188-202). London: Continuum.
- Fellbaum, C. (2007b). Introduction. In C. Fellbaum (Hrsg.), *Idioms and collocations* (S. 1-19). London: Continuum.
- Fellbaum, C., Geyken, A., Herold, A., Koerner, F. & Neumann, G. (2006). Corpus-based Studies of German Idioms and Light Verbs. *International Journal of Lexicography*, 19 (4), 349-360.
- Ferguson, C. A. (1994). Dialect, Register, and Genre: Working Assumptions About Conventionalization. In D. Biber & E. Finegan (Hrsg.), *Sociolinguistic perspectives on register* (S. 15-30). New York: Oxford University Press.
- Fillmore, C. J. (1968). The case for case. In E. Bach & R. Harms (Hrsg.), *Universals in Linguistic Theory*. New York: Holt, Winehart and Winston.
- Fillmore, C. J. (Hrsg.). (2003). *Form and Meaning in Language: Peppers on semantic roles* (Bd. 1). Leland, CA: CSLI Publications.
- Fillmore, C. J., Johnson, C. R. & Petruck, M. R. L. (2003). Background to Framenet. *International Journal of Lexicography*, 16 (3), 235-250.
- Fillmore, C. J., Kay, P. & O'Connor, M. C. (1988). Regularity and idiomaticity

- in grammatical constructions: The case of let alone. *Language*, 64 (3), 510-538.
- Fischer, K. (2007). Konstruktionsgrammatik und Interaktion. In K. Fischer & A. Stefanowitsch (Hrsg.), *Konstruktionsgrammatik* (Bd. 40, S. 133-150). Tübingen: Stauffenburg-Verlag.
- Fischer, K. & Stefanowitsch, A. (2007a). Konstruktionsgrammatik: Ein Überblick. In K. Fischer & A. Stefanowitsch (Hrsg.), *Konstruktionsgrammatik* (Bd. 40, S. 3-17). Tübingen: Stauffenburg-Verlag.
- Fischer, K. & Stefanowitsch, A. (Hrsg.). (2007b). *Konstruktionsgrammatik: Von der Anwendung zur Theorie* (Bd. 40). Tübingen: Stauffenburg-Verlag.
- Fontao, M. I. & Mergenthaler, E. (2008). Therapeutic factors and language patterns in group therapy application of computer-assisted text analysis to the examination of microprocesses in group therapy: Preliminary findings. *Psychotherapy Research*, 18 (3), 345-354.
- Ford, C. E., Fox, B. A. & Thompson, S. A. (Hrsg.). (2002). *The language of turn and sequence*. Oxford: Oxford University Press.
- Ford, C. E., Fox, B. A. & Thompson, S. A. (2003). Social Interaction and Grammar. In M. Tomasello (Hrsg.), *The new psychology of language* (Bd. 2, S. 119-144). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Ford, C. E. & Thompson, S. A. (1996). Interactional units in conversation: Syntactic, intonational, and pragmatic resources for the management of turns. In E. Ochs, E. A. Schegloff & S. A. Thompson (Hrsg.), *Interaction and grammar* (Bd. 13, S. 134-184). Cambridge etc.: Cambridge University Press.
- Franceschini, R. (2006). Bedeutungskonstitution mittels sequenzieller Erwartungen am Beispiel des "Quasi-Italienisch" von Deutschsprachigen. In A. Depermann & T. Spranz-Fogasy (Hrsg.), *Be-deuten* (S. 119-133). Tübingen: Stauffenburg.
- Freud, S. (1904). *Zur Psychopathologie des Alltagslebens*. GW IV..
- Fried, M. & Östman, J.-O. (2005). Construction Grammar and spoken language: The case of pragmatic particles: Approaches to Spoken Interaction. *Journal of Pragmatics*, 37 (11), 1752-1778.
- Georgaca, E. & Avdi, E. (2009). Evaluating the Talking Cure: The Contribution of Narrative, Discourse, and Conversation Analysis to Psychotherapy Assessment. *Qualitative Research in Psychology*, 6 (3), 233-247.
- Georgakopoulou, A. (2007). *Small stories, interaction and identities* (Bd. 8). Amsterdam: John Benjamins Publishing.
- Gibbs, R. W. (1989). Understanding and literal meaning. *Cognitive Science*, 13 (2), 243-251.
- Gibbs, R. W. (1994). *The poetics of mind. Figurative thought, language and understanding* (2. Aufl., Bd. 2). Cambridge: Cambridge University Press.
- Gibbs, R. W. (2002). A new look at literal meaning in understanding what is said and implicated: Literal, minimal, and salient meanings. *Journal of*

- Pragmatics*, 34 (4), 457–486.
- Glück, H. (Hrsg.). (2005). *Metzler-Lexikon Sprache* (3. Aufl.). Stuttgart: Metzler.
- Goffman, E. (2008). *Wir alle spielen Theater: Die Selbstdarstellung im Alltag* (6. Aufl. Aufl., Bd. 3891). München: Piper.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: Univ. of Chicago Press.
- Götz-Votteler, K. (2007). Describing semantic valency. In T. Herbst & K. Götz-Votteler (Hrsg.), *Valency* (Bd. 187, S. 37-50). Berlin: Mouton de Gruyter.
- Granger, S. & Meunier, F. (2008). Introduction: The many faces of phraseology. In S. Granger & F. Meunier (Hrsg.), *Phraseology* (S. XX–XXVIII). Amsterdam: John Benjamins.
- Granger, S. & Paquot, M. (2008). Disentangling the phraseological web. In S. Granger & F. Meunier (Hrsg.), *Phraseology* (S. 27-50). Amsterdam: John Benjamins.
- Grégoire, N. (2006). Elaborating the parameterized Equivalence Class Method for Dutch. *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC 2006)*, 1894–1899.
- Gries, S. T. (2008). Phraseology and linguistic theory: A brief survey. In S. Granger & F. Meunier (Hrsg.), *Phraseology* (S. 3-25). Amsterdam: John Benjamins.
- Grimmer, B. (2000). Kreditierung in einer psychoanalytisch orientierten Psychotherapie. *Psychotherapie und Sozialwissenschaft*, 2 (4).
- Gülich, E. (2001). Zum Zusammenhang von alltagsweltlichen und wissenschaftlichen "Methoden". In K. Brinker, G. Antos, W. Heinemann & S. F. Sager (Hrsg.), *Text- und Gesprächslinguistik* (Bd. 2. Halbband, S. 1086–1092). Berlin: de Gruyter.
- Gülich, E. (2007a). Mündliches Erzählen: narrative und szenische Rekonstruktion. In U. Weise, A. Ruwe, L. Jonker & S. Lubs (Hrsg.), *Behutsames Lesen* (S. 35–55). Evangelische Verlagsanstalt.
- Gülich, E. (2007b). "Volle Palette in Flammen". Zur Orientierung an vorgeformten Strukturen beim Reden über Angst. *Psychotherapie und Sozialwissenschaft*, 9 (1).
- Günthner, S. (2005). Narrative reconstructions of past experiences: Adjustment and modifications in the process of recontextualizing a past experience. In U. M. Quasthoff & T. Becker (Hrsg.), *Narrative interaction* (Bd. 5, S. 285–301). Amsterdam: John Benjamins Publishing.
- Günthner, S. (2006a). Rhetorische Verfahren bei der Vermittlung von Panikattacken: Zur Kommunikation von Angst in informellen Gesprächskontexten. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion*, 7, 124–151.
- Günthner, S. (2006b). Von Konstruktionen zu kommunikativen Gattungen: Die Relevanz sedimentierter Muster für die Ausführung kommunikativer Aufgaben. *gidi Arbeitspapierreihe*, 1, 1–21.

- Günthner, S. (2007a). Brauchen wir eine Theorie der gesprochenen Sprache? Und: wie kann sie aussehen? - Ein Plädoyer für eine praxisorientierte Grammatiktheorie. *gidi Arbeitspapierreihe*, 6, 1–22.
- Günthner, S. (2007b). Zur Emergenz grammatischer Funktionen im Diskurs: wo-Konstruktionen in Alltagsinteraktionen. In H. Hausendorf (Hrsg.), *Gespräch als Prozess* (Bd. 37, S. 125–155). Tübingen: Gunter Narr.
- Günthner, S. (2008). Eine Grammatik der Theatralität? Grammatische und prosodische Inszenierungsverfahren in Alltagserzählungen. *gidi Arbeitspapierreihe*, 2, 1–27.
- Günthner, S. & Bucker, J. (2009). Einleitung zum Sammelband. In S. Günthner & J. Bucker (Hrsg.), *Grammatik im Gespräch* (S. 1–22). de Gruyter.
- Günthner, S. & Imo, W. (Hrsg.). (2006a). *Konstruktionen in der Interaktion* (Bd. 20). Berlin: de Gruyter.
- Günthner, S. & Imo, W. (2006b). Konstruktionen in der Interaktion. In S. Günthner & W. Imo (Hrsg.), *Konstruktionen in der Interaktion* (Bd. 20, S. 1–22). Berlin: de Gruyter.
- Haase, M. (2010). *Szenen einer Ehe – Erzählungen des Gustav Y.* Unveröffentlichte Lizentiatsarbeit, Universität Zürich, Psychologisches Institut, Klinische Psychologie, Psychotherapie und Psychoanalyse.
- Häcki Buhofer, A. (2004). Spielräume des sprachverstehens: Psycholinguistische zugänge zum individuellen umgang mit phraseologismen. In K. Steyer (Hrsg.), *Wortverbindungen - mehr oder weniger fest* (Bd. 2003, S. 144–164). Berlin: de Gruyter.
- Hakulinen, A. & Selting, M. (2005a). Introduction. In A. Hakulinen & M. Selting (Hrsg.), *Syntax and lexis in conversation* (Bd. Vol. 17, S. 1–13). Amsterdam: John Benjamins.
- Hakulinen, A. & Selting, M. (Hrsg.). (2005b). *Syntax and lexis in conversation: Studies on the use of linguistic resources in talk-in-interaction* (Bd. Vol. 17). Amsterdam: John Benjamins.
- Halliday, M. A. K. & Matthiessen, C. M. (2004). *An introduction to functional grammar* (3rd ed. Aufl.). London: Arnold.
- Hanks, P. (2004). The Syntagmatics of Metaphor and Idiom. *International Journal of Lexicography*, 17 (3), 245–274.
- Hanks, P. (2005). Similes and sets: the english preposition like. In R. Blatná & V. Petkevič (Hrsg.), *Jazyky a jazykoveda (languages and linguistics: Festschrift for professor fr. cermak)*. Prague: Prague: Philosophy Faculty of the Charles University.
- Hanks, P. (2008a). Lexical Patterns: From Hornby to Hunston and beyond. In E. Bernal & J. DeCesaris (Hrsg.), *Proceedings of the XIII. Euralex International Congress* (S. 89–129).
- Hanks, P. (2008b). The Lexicographical Legacy of John Sinclair. *Int J Lexicography*, 219–229.
- Hanks, P. (2010). How people use words to make meanings. *NLPCS 2010*

- Proceedings*, 1–11.
- Hanks, P. (in press). *Analyzing the Lexicon: Norms and Exploitations*. MIT Press.
- Hanks, P. & Jezek, E. (2008). Shimmering Lexical Sets. In E. Bernal & J. De-Cesaris (Hrsg.), *Proceedings of the XIII. Euralex International Congress* (S. 391-402).
- Hanks, P. & Pustejovsky, J. (2005). A pattern dictionary for natural language processing. *Revue Francaise de Langue Appliquée*, 10 (2), 1–19.
- Hanks, P., Urbschat, A. & Gehweiler, E. (2006). German Light Verb Constructions in Corpora and Dictionaries. *Int J Lexicography*, 19 (4), 439-457.
- Hardtke, K., Levitt, H. & Angus, L. E. (2003). Narrative Prozesse im Beratungs- und Psychotherapiediskurs: Das Narrative Processes Coding System (NP-CS). *Zeitschrift für Qualitative Sozialforschung* (1).
- Harras, G., Proost, K. & Winkler, E. (2007). *Handbuch deutscher Kommunikationsverben: Teil 2: Lexikalische Strukturen* (Bde. 10,2). Berlin: de Gruyter.
- Harras, G., Winkler, E., Erb, S. & Proost, K. (2004). *Handbuch deutscher Kommunikationsverben* (Bde. 10,1). Berlin: de Gruyter.
- Haspelmath, M. (2002). Grammatikalisierung: von der Performanz zur Kompetenz ohne angeborene Grammatik. In S. Krämer & E. König (Hrsg.), *Gibt es eine Sprache hinter dem Sprechen?* (Bd. 1592, S. 1-19). Frankfurt am Main: Suhrkamp.
- Hausmann, F. J. (2004). Was sind eigentlich kollokationen? In K. Steyer (Hrsg.), *Wortverbindungen - mehr oder weniger fest* (Bd. 2003, S. 309–334). Berlin: de Gruyter.
- Heid, U. (2007). Computational linguistic aspects of phraseology II. In H. Burger, D. Dobrovolskij, P. Kühn & N. R. Norrick (Hrsg.), *Phraseologie* (S. 1036-1044). Berlin: W. de Gruyter.
- Heid, U. (2008). Computational phraseology: An overview. In S. Granger & F. Meunier (Hrsg.), *Phraseology* (S. 337-360). Amsterdam: John Benjamins.
- Helbig, G. (1992). *Probleme der Valenz- und Kasus-theorie: Konzepte der Sprach- und Literaturwissenschaft 51*. Tübingen: Niemeyer.
- Helbig, G. & Buscha, J. (2000). *Leitfaden der deutschen Grammatik*. Berlin: Langenscheidt.
- Herman, D. (2003). Introduction. In D. Herman (Hrsg.), *Narrative theory and the cognitive sciences* (Bd. no. 158, S. 1–30). Stanford, Calif.: CSLI.
- Herman, D., Jahn, M. & Ryan, M.-L. (Hrsg.). (2005). *Routledge encyclopedia of narrative theory*. London: Routledge.
- Hoey, M. (2005). *Lexical priming: A new theory of words and language*. London: Routledge.
- Hölzer, M., Scheytt, N. & Kächele, H. (1992). Das "Affektive Diktionär Ulm" als eine Methode der quantitativen Vokabularbestimmung. In C. Züll &

- P. P. Mohler (Hrsg.), *Textanalyse* (S. 131–154). Opladen: Westdeutscher Verlag.
- Hopper, P. J. (1987). Emergent Grammar. *Berkeley Linguistics Society*, 13, 139–157.
- Hopper, P. J. & Thompson, S. A. (1984). The Discourse Basis for Lexical Categories in Universal Grammar. *Language*, 60 (4), 703–752.
- Hörmann, H. (1988). *Meinen und Verstehen: Grundzüge einer psychologischen Semantik* (3. Aufl. Aufl., Bd. 230). Frankfurt Main: Suhrkamp.
- Imo, W. (2007). *Construction Grammar und Gesprochene-Sprache-Forschung: Konstruktionen mit zehn matrixsatzfähigen Verben im gesprochenen Deutsch* (Bd. 275). Tübingen: Niemeyer.
- Ivanova, K., Heid, U., Schulte Im Walde, S., Kilgariff, A. & Pomikalek, J. (2008). Evaluating a German Sketch Grammar: A Case Study on Noun Phrase Case. *Proceedings LREC. Marrakech, Morocco*, 2101–2107.
- Jezek, E. & Hanks, P. (2010). What lexical sets tell us about conceptual categories. *Lexis - E-Journal in English Lexicology*, 4: *Corpus Linguistics and the Lexicon*, 7–22.
- Jung, E. (2009). *Beziehung ist Krieg, Therapie ist Weg und ich bin ein Stern. Eine Systematische Metaphernanalyse anhand von transkribierten Therapiestunden der Frau W.* Unveröffentlichte Lizentiatsarbeit, Universität Zürich, Psychologisches Institut, Klinische Psychologie, Psychotherapie und Psychoanalyse.
- Kallmeyer, W. (1985). Handlungskonstitution im Gespräch. In E. Gülich & T. Kotschi (Hrsg.), *Grammatik, Konversation, Interaktion* (Bd. 153, S. 81–123). Tübingen: Niemeyer.
- Kamber, A. (2006). Funktionsverbgefüge - empirisch (am Beispiel von 'kommen'). *Linguistik Online*, 28 (3), 109–132.
- Kamber, A. (2008). *Funktionsverbgefüge - empirisch: Eine korpusbasierte Untersuchung zu den nominalen Prädikaten des Deutschen* (Bd. 281). Tübingen: Niemeyer.
- Keil, M. (1997). *Wort für wort: Repräsentation und verarbeitung verbaler phraseologismen (phraseo-lex): Diss., univ. erlangen-nürnberg* (Bd. 35). Tübingen: Niemeyer.
- Kern, F. & Quasthoff, U. M. (2005). Fantasy Stories and conversational narratives of personal experience: Genre-specific, interactional and developmental perspectives. In U. M. Quasthoff & T. Becker (Hrsg.), *Narrative interaction* (Bd. 5, S. 15–56). Amsterdam: John Benjamins Publishing.
- Kilgariff, A. (2008). I don't believe in word senses. In T. Fontenelle (Hrsg.), *Practical lexicography* (S. 135–152). Oxford: Oxford University Press.
- Kilgariff, A., Rychlý, P., Smrz, P. & Tugwell, D. (2004). The Sketch Engine. *Proc EURALEX 2004, Lorient, France*.
- Klein, C. & Martinez, M. (Hrsg.). (2009). *Wirklichkeitserzählungen: Felder, Formen und Funktionen nicht-literarischen Erzählens*. Stuttgart: Metzler.

- Konerding, K.-P. (1993). *Frames und lexikalisches Bedeutungswissen. Untersuchungen zur linguistischen Grundlegung einer Frametheorie und zu ihrer Anwendung in der Lexikographie*. Tübingen: Niemeyer.
- Kuiper, K. (2009). *Formulaic genres*. Basingstoke: Palgrave Macmillan.
- Kunze, C. & Lemnitzer, L. (2007). *Computerlexikographie: Eine Einführung*. Tübingen: Narr.
- Kurri, K. & Wahlström, J. (2007). Reformulations of agentless talk in psychotherapy. *Text & Talk - An Interdisciplinary Journal of Language, Discourse & Communication Studies*, 27 (3), 315–338.
- Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia Pa.: University of Pennsylvania Press.
- Labov, W. & Waletzky, J. (1967). Narrative Analysis: Oral Versions of Personal Experience. In J. Helm (Hrsg.), *Essays on the verbal and visual arts* (S. 12–44). Seattle: University of Washington Press.
- Lakoff, G. (1987). *Women, Fire and Dangerous Things: What Categories Reveal About the Mind*. Chicago: The University of Chicago Press.
- Lakoff, G. (2009). *The Political Mind: A Cognitive Scientist's Guide to Your Brain and Its Politics*. Penguin Books.
- Lakoff, G. & Johnson, M. (1980). *Metaphors we live by*. Chicago: University of Chicago Press.
- Langacker, R. W. (1987). *Foundations of cognitive grammar*. Stanford, CA: Stanford University Press.
- Langacker, R. W. (1999). *Grammar and conceptualization* (Bd. 14). Berlin: Mouton de Gruyter.
- Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.
- Langacker, R. W. (2009). Cognitive (Construction) Grammar. *Cognitive Linguistics*, 20 (1), 167–176.
- Langer, S. (2005). A Formal Specification of Support Verb Constructions. In S. Langer & D. Schnorbusch (Hrsg.), *Semantik im Lexikon* (S. 179–201). Tübingen: Narr.
- Lee, D. Y. W. (2001). Genres, Registers, Text types, Domains, and Styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language Learning & Technology*, 5 (3), 37–72.
- Lemnitzer, L. & Zinsmeister, H. (2006). *Korpuslinguistik: Eine Einführung*. Tübingen: Gunter Narr Verlag.
- Lepper, G. & Mergenthaler, E. (2007). Observing therapeutic interaction in the "Lisa" case. *Psychotherapy Research*, 17, 1–11.
- Levin, B. (1993). *English Verb Classes and Alternations*. Chicago: University of Chicago Press.
- Linell, P. (1998). *Approaching dialogue: Talk, interaction and contexts in dialogical perspectives* (Bd. vol. 3). Amsterdam: Benjamins.
- Linell, P. (2009). *Rethinking language, mind, and world dialogically: Interactional*

- and contextual theories of human sense-making*. Charlotte, N.C.: Information Age Publishing.
- Luborsky, L. & Crits-Christoph, P. (1990). *Understanding Transference: Core Conflictual Relationship Theme Method*. The Perseus Books Group.
- Lucius-Hoene, G. & Deppermann, A. (2004). Narrative Identität und Positionierung. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion*, 5, 166–183. Verfügbar unter www.gespraechsforschung-ozs.de
- Luder, M. (1999). *Kategorien und Codes: Auf dem Weg zu einer computergestützten Fassung der Erzählanalyse JAKOB*. Unveröffentlichte Dissertation, Universität Zürich, Psychologisches Institut, Klinische Psychologie I.
- Luder, M. (2009). Konstruktionen in der Erzählanalyse JAKOB. In G. Grimm, N. Kapfhamer, H. Mathys, S. Michel & B. Boothe (Hrsg.), *Erzählen, Träumen, Erinnern. Erträge Klinischer Erzählforschung* (S. 226–237). Lengerich: Pabst.
- Luder, M. (2010). Building an OLIF-based lexical database for representing constructions. In S. Granger & M. Paquot (Hrsg.), *eLexicography in the 21st century: New challenges, new applications* (Bd. 7, S. 229–239). Louvain-la-Neuve: Presses universitaires de Louvain.
- Luder, M. & Clematide, S. (2010). Constructing a Constructional MWE Lexicon for psycho-conceptual Annotation: Evaluation of CPA and DuELME for Lexicographic Description. In A. Dykstra & T. Schoonheim (Hrsg.), *Proceedings of the XIV Euralex International Congress* (S. 152–153).
- Luder, M., Clematide, S. & Distl, B. (2008). Ein elektronisches Lexikon im OLIF-Format für die Erzählanalyse. In E. Bernal & J. DeCesaris (Hrsg.), *Proceedings of the XIII. Euralex International Congress* (S. 729–735).
- Luif, V. (2006). Die Psychose als Erzählgeschehen. Strukturelle, inhaltliche und rezeptionsorientierte Merkmale der Erzählungen aus dem Tagebuch eines Schizophrenen. In V. Luif, G. Thoma & B. Boothe (Hrsg.), *Beschreiben - Erschliessen - Erläutern. Psychotherapieforschung als qualitative Wissenschaft* (S. 66–89). Lengerich: Pabst.
- MacLagan, M., Davis, B. & Lunsford, R. (2008). Fixed expressions, extenders and metonymy in the speech of people with Alzheimer's disease. In S. Granger & F. Meunier (Hrsg.), *Phraseology* (S. 175–187). Amsterdam: John Benjamins.
- McCormick, S. M. (2005). *The Structure and Content of the Body of an OLIF v.2.0/2.1 File*. Verfügbar unter <http://www.olif.net/>
- McCormick, S. M., Lieske, C. & Culum, A. (2004). *OLIF v.2: A Flexible Language Data Standard*. Verfügbar unter <http://www.olif.net/>
- McLeod, J. (1997). *Narrative and psychotherapy*. London: Sage.
- Mergenthaler, E. (1992). *Die Transkription von Gesprächen* (3. Aufl.). Ulm: Ulmer Textbank.
- Mergenthaler, E. (1996). Emotion-abstraction patterns in verbatim protocols: a new way of describing psychotherapeutic processes. *Journal Of Consulting*

- And Clinical Psychology*, 64 (6), 1306–1315.
- Mergenthaler, E. (1997). *Emotions/Abstraktionsmuster in Verbatimprotokollen*. Frankfurt: VAS Verlag für Akademische Schriften.
- Mergenthaler, E. & Kächele, H. (1994). Die Ulmer Textbank. *Psychotherapie Psychosomatik Medizinische Psychologie*, 44 (1), 29–35.
- Meyer, K. & Bickel, H. (2006). *Schweizer Wörterbuch: So sagen wir in der Schweiz*. Frauenfeld: Huber Frauenfeld.
- Moon, R. (1998). *Fixed expressions and idioms in english: A corpus-based approach*. Oxford: Clarendon Press.
- Moon, R. (2008). Sinclair, Phraseology, and Lexicography. *International Journal of Lexicography*, 21 (3), 243–254.
- Müller, S. (2006). Phrasal or Lexical Constructions? *Language*, 82 (4), 850–883.
- Müller, S. (2007). Resultativkonstruktionen, Partikelverben und syntaktische vs. lexikonbasierte Konstruktionen. In K. Fischer & A. Stefanowitsch (Hrsg.), *Konstruktionsgrammatik* (Bd. 40). Tübingen: Stauffenburg-Verlag.
- Müller, S. (2008). *Head-driven phrase structure grammar: Eine Einführung* (2. Aufl., Bd. 17). Tübingen: Stauffenburg.
- Murphy, M. L., Paradis, C. & Willners, C. (2009). Introduction: Lexical contrast in discourse: Lexical Contrast in Discourse. *Journal of Pragmatics*, 41 (11), 2137–2139.
- Norricks, N. R. (2001). Discourse markers in oral narrative. *Journal of Pragmatics*, 33 (6), 849–878.
- Ochs, E. & Capps, L. (2001). *Living narrative: Creating lives in everyday storytelling*. Cambridge Mass. u.a.: Harvard Univ. Press.
- Ochs, E., Schegloff, E. A. & Thompson, S. A. (Hrsg.). (1996). *Interaction and grammar* (Bd. 13). Cambridge etc.: Cambridge University Press.
- Ono, T. & Thompson, S. A. (1995). What can Conversation Tell us about Syntax? In P. W. Davis (Hrsg.), *Alternative linguistics* (Bd. 102, S. 213–271). Amsterdam etc.: Benjamins.
- Östman, J.-O. (2005). Construction Discourse: A prolegomenon. In J.-O. Östman & M. Fried (Hrsg.), *Construction grammars* (Bd. 3, S. 121–144). Amsterdam: Benjamins.
- Östman, J.-O. & Fried, M. (2005a). The cognitive grounding of Construction Grammar. In J.-O. Östman & M. Fried (Hrsg.), *Construction grammars* (Bd. 3, S. 1–13). Amsterdam: Benjamins.
- Östman, J.-O. & Fried, M. (Hrsg.). (2005b). *Construction grammars: Cognitive grounding and theoretical extensions* (Bd. 3). Amsterdam: Benjamins.
- Overlach, F. (2008). *Sprache des Schmerzes - Sprechen über Schmerzen: Eine grammatisch-semantische und gesprächsanalytische Untersuchung von Schmerzausdrücken im Deutschen* (Bd. 30). Berlin: de Gruyter.
- Panther, K.-U. & Radden, G. (2005). Metonymy. In J.-O. Östman & J. Verschueren (Hrsg.), *Handbook of pragmatics* (S. 1–14). Amsterdam: Benjamins.

- Partington, A. (1998). *Patterns and meanings: using corpora for english language research and teaching*. J. Benjamins Pub.
- Pascual, C. Perinán, Arcas-Túnez, F., Florencio, E., Amores, G., Pérez, G., Manchón, P. et al. (2008). Modelling OLIF frame with EAGLES/ISLE specifications: an interlingual approach. *Procesamiento del Lenguaje Natural*, 40, 9-16.
- Pawley, A. & Syder, F. H. (1983). Two Puzzles for Linguistic Theory: Native-like Selection and Native-like Fluency. In J. C. Richards & R. W. Schmidt (Hrsg.), *Language and Communication* (S. 191-226). New York: Longman.
- Pedrazzini, S. (1994). *Phrase manager: A system for phrasal and idiomatic dictionaries* (Bd. 3). Hildesheim: Olms.
- Pennebaker, J. W. & Francis, M. E. (2001). *Linguistic Inquiry and Word Count*. Hillsdale (N.J.): Lawrence Erlbaum Associates.
- Pennebaker, J. W., Mehl, M. R. & Niederhoffer, K. G. (2003). Psychological Aspects of Natural Language Use: Our Words, Our Selves. *Annual Review of Psychology*, 54, 547-577.
- Peräkylä, A. (2004). Making Links in Psychoanalytic Interpretations: A Conversation Analytical Perspective. *Psychotherapy Research*, 14 (3), 289-307.
- Polenz, P. v. (2008). *Deutsche Satzsemantik: Grundbegriffe des Zwischen-den-Zeilen-Lesens* (3. Aufl.). Berlin: de Gruyter.
- Pottelberge, J. van. (2001). *Verbonominale Konstruktionen, Funktionsverbgefüge: Vom Sinn und Unsinn eines Untersuchungsgegenstandes* (Bd. 12). Heidelberg: Winter.
- Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, Mass.: MIT Press.
- Pustejovsky, J., Hanks, P. & Rumshisky, A. (2004). Automated induction of sense in context. *COLING 2004, Geneva*, 1-7.
- Radden, G., Köpcke, K.-M., Berg, T. & Siemund, P. (2007). Introduction: The construction of meaning in language. In G. Radden, K.-M. Köpcke, T. Berg & P. Siemund (Hrsg.), *Aspects of meaning construction* (S. 1-18). Amsterdam: Benjamins.
- Rieser, M. (2010). *Im Schatten des Vaters - Erzählungen des Gustav Y. Eine narrative Einzelfallstudie der Psychodynamik der Vater-Sohn-Beziehung auf der Basis der Erzählanalyse JAKOB*. Unveröffentlichte Lizentiatsarbeit, Universität Zürich, Psychologisches Institut, Klinische Psychologie, Psychotherapie und Psychoanalyse.
- Roth, S. (2006). *Lexikalisch-semantische Netze: Anwendungsperspektiven für die Computerlinguistik und Sprachtechnologie: Eine Einführung*. Unveröffentlichte Dissertation, Universität Zürich, Philosophische Fakultät, Institut für Computerlinguistik.
- Rotlkegel, A. (2007). Computerlinguistische aspekte der phraseme i. In H. Burger, D. Dobrovolskij, P. Kühn & N. R. Norrick (Hrsg.), *Phraseologie* (S. 1027-1035). Berlin: W. de Gruyter.
- Rullkoetter, N., Bullig, R., Driessen, M., Beblo, T., Mensebach, C. & Wingenfeld,

- K. (2008). Autobiographical Memory and Language Use: Linguistic Analyses of Critical Life Event Narratives in a Non-Clinical Population. *Applied Cognitive Psychology*, 23, 278–287.
- Ruppenhofer, J., Ellsworth, M., Petruck, M. R. L., Johnson, C. R. & Scheffczyk, J. (2010). *FrameNet II: Extended Theory and Practice*. Berkeley. Verfügbar unter <http://framenet.icsi.berkeley.edu/>
- Rychlý, P. (2008). A lexicographer-friendly Association Score. In P. Sojka & A. Horák (Hrsg.), *Proceedings of Recent Advances in Slavonic Natural Language Processing* (S. 6–9). Brno: Masaryk University.
- Sager, S. F. (2001). Gesprächssorte - Gesprächstyp - Gesprächsmuster. In K. Brinker, G. Antos, W. Heinemann & S. F. Sager (Hrsg.), *Text- und Gesprächslinguistik* (Bd. 2. Halbband, S. 1464–1471). Berlin: de Gruyter.
- Sarbin, T. R. (Hrsg.). (1986). *Narrative Psychology: The Storied Nature of Human Conduct*. Greenwood Press.
- Schank, R. C. & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Hillsdale, New Jersey: John Wiley and Sons.
- Schegloff, E. A., Ochs, E. & Thompson, S. A. (1996). Introduction. In E. Ochs, E. A. Schegloff & S. A. Thompson (Hrsg.), *Interaction and grammar* (Bd. 13, S. 1–51). Cambridge etc.: Cambridge University Press.
- Schmitt, R. (2010). Metaphernanalyse. In G. Mey & K. Mruck (Hrsg.), *Handbuch qualitative forschung in der psychologie* (S. 676–691). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schmitt, R. (in press). *Einführung in die systematische metaphernanalyse*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Schneider, G. (2008). *Hybrid Long-Distance Functional Dependency Parsing*. Unveröffentlichte Dissertation, Institute of Computational Linguistics, University of Zurich.
- Schnorbusch, D. (2005). Semantische Klassen aus syntaktischen Klassen? In S. Langer & D. Schnorbusch (Hrsg.), *Semantik im Lexikon* (S. 33–58). Tübingen: Narr.
- Scholz-Stubenrecht, W. & Wermke, M. (2008). *Duden 11 - Redewendungen: Wörterbuch der deutschen Idiomatik* (3. Aufl., Bd. 3). Mannheim: Dudenverlag.
- Schulte Im Walde, S. (2005). Induction of semantic classes for german verbs. In S. Langer & D. Schnorbusch (Hrsg.), *Semantik im Lexikon* (S. 59–86). Tübingen: Narr.
- Schumacher, H., Kubczak, J., Schmidt, R. & Ruiter, V. de. (2004). *VALBU - Valenzwörterbuch deutscher Verben*. Tübingen: Narr.
- Schütze, F. (1982). Narrative Repräsentation kollektiver Schicksalsbetroffenheit. *Erzählforschung*, 7, 568–590.
- Schütze, F. (1983). Biographieforschung und narratives Interview. *Neue Praxis*, 13 (3), 283–293.

- Schwitalla, J. (2001). Gesprochene-Sprache-Forschung und ihre Entwicklung zu einer Gesprächsanalyse. In K. Brinker, G. Antos, W. Heinemann & S. F. Sager (Hrsg.), *Text- und Gesprächslinguistik* (Bd. 2. Halbband, S. 896–902). Berlin: de Gruyter.
- Schwitalla, J. (2003). *Gesprochenes Deutsch: Eine Einführung* (2., überarb. Aufl. Aufl.). Berlin: Erich Schmidt Verlag.
- Schwitalla, J. & Tiittula, L. (2009). *Mündlichkeit in literarischen Erzählungen: Sprach- und Dialoggestaltung in modernen deutschen und finnischen Romanen und deren Übersetzung*. Stauffenburg.
- Selting, M., Auer, P., Barden, B., Bergmann, J., Couper-Kuhlen, E., Günthner, S. et al. (1998). Gesprächsanalytisches Transkriptionssystem (GAT). *Linguistische Berichte* (173), 99–122.
- Selting, M. & Couper-Kuhlen, E. (2000). Argumente für die Entwicklung einer 'interaktionalen Linguistik'. *Gesprächsforschung - Online-Zeitschrift zur verbalen Interaktion* (1), 76–95.
- Selting, M. & Couper-Kuhlen, E. (Hrsg.). (2001). *Studies in interactional linguistics* (Bd. 10). Amsterdam: John Benjamins.
- Sennrich, R., Schneider, G., Volk, M. & Warin, M. (2009). A New Hybrid Dependency Parser for German. In C. Chiarcos, R. E. d. Castilho & M. Stede (Hrsg.), *Von der Form zur Bedeutung: Texte automatisch verarbeiten* (S. 115–124). Tübingen: Gunter Narr.
- Shadrach Long, P. & Lepper, G. (2008). Metaphor in psychoanalytic psychotherapy: a comparative study of four cases by a practitioner-researcher. *British Journal of Psychotherapy*, 24 (3), 343–364.
- Siepmann, D. (2006). Review of Fischer, Kerstin & Anatol Stefanowitsch (eds.) 2006. "Konstruktionsgrammatik. Von der Anwendung zur Theorie". *constructions-online*. Verfügbar unter <http://www.constructions-online.de/reviews>
- Sinclair, J. (1991). *Corpus, concordance, collocation: Describing english language*. Oxford University Press.
- Smirnova, E. & Mortelmans, T. (2010). *Funktionale Grammatik: Konzepte und Theorien*. Berlin: de Gruyter.
- Söhn, J.-P. (2006). *Über bärendienste und erstaunte bauklötze: Idiome ohne freie lesart in der hpsg* (Bd. 1930). Frankfurt am Main: P. Lang.
- Söhn, J.-P. & Römer, C. (2004). Wann ist ein idiom ein idiom? eine analyse von phraseologismen ohne freie lesart. *Phraseology in Motion, Proceedings zu Europhras 2004* (Basel), 161–172.
- Spence, D. P. (1980). Lawfulness in lexical choice: a natural experiment. *Journal of the American Psychoanalytic Association*, 28, 115–132.
- Spence, D. P. (1984). *Narrative Truth and Historical Truth: Meaning and Interpretation in Psychoanalysis*. New York: Norton.
- Stärk, F. (2008). *Geschwisterbeziehung in den Erzählungen von Frau W. Eine narrative Einzelfallstudie*. Unveröffentlichte Lizentiatsarbeit, Universität

- Zürich, Psychologisches Institut, Klinische Psychologie, Psychotherapie und Psychoanalyse.
- Stärk, F. (2009). Geschwisterbeziehung in den Erzählungen von Frau W. - Eine narrative Einzelfallstudie der Psychodynamik in einer Geschwisterbeziehung auf der Basis der Erzählanalyse JAKOB. In G. Grimm, N. Kapfhamer, H. Mathys, S. Michel & B. Boothe (Hrsg.), *Erzählen, Träumen, Erinnern. Erträge Klinischer Erzählforschung* (S. 209-217). Lengerich: Pabst.
- Stefanowitsch, A. & Fischer, K. (Hrsg.). (2008). *Konstruktionsgrammatik II: Von der Konstruktion zur Grammatik* (Bd. 47). Tübingen: Stauffenburg.
- Stein, S. (2004). Formelhaftigkeit und Routinen in mündlicher Kommunikation. In K. Steyer (Hrsg.), *Wortverbindungen - mehr oder weniger fest* (Bd. 2003, S. 262-288). de Gruyter.
- Steyer, K. (2004a). Kookkurrenz. Korpusmethodik, linguistisches Modell, lexicographische Perspektiven. In K. Steyer (Hrsg.), *Wortverbindungen - mehr oder weniger fest* (Bd. 2003, S. 87-116). de Gruyter.
- Steyer, K. (Hrsg.). (2004b). *Wortverbindungen - mehr oder weniger fest* (Bd. 2003). Berlin: de Gruyter.
- Steyer, K. & Brunner, A. (2009). Das UWV-Analysemodell. Eine korpusgesteuerte Methode zur linguistischen Systematisierung von Wortverbindungen. *OPAL - Online publizierte Arbeiten zur Linguistik* (1).
- Storjohann, P. (2009). Plesionymy: A case of synonymy or contrast? Lexical Contrast in Discourse. *Journal of Pragmatics*, 41 (11), 2140-2158.
- Storrer, A. (2007). Corpus-based investigations on German support verb constructions. In C. Fellbaum (Hrsg.), *Idioms and collocations* (S. 164-187). London: Continuum.
- Streeck, U. (2008). A psychotherapist's view of conversation analysis. In A. Peräkylä, C. Antaki, S. Vehviläinen & I. Leudar (Hrsg.), *Conversation analysis and psychotherapy* (S. 173-187). Cambridge: Cambridge University Press.
- ten Have, P. (2007). *Doing conversation analysis: A practical guide* (2. Aufl.). London: Sage.
- Thomä, H. & Kächele, H. (2006). *Psychoanalytische Therapie: Grundlagen* (3., überarb. und aktualisierte Aufl. Aufl., Bd. 1). Heidelberg: Springer.
- Tomasello, M. (2003). Introduction: Some surprises for psychologists. In M. Tomasello (Hrsg.), *The new psychology of language* (Bd. 2, S. 1-14). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Tomasello, M. (2007). Konstruktionsgrammatik und früher Erstspracherwerb. In K. Fischer & A. Stefanowitsch (Hrsg.), *Konstruktionsgrammatik* (Bd. 40, S. 19-37). Tübingen: Stauffenburg-Verlag.
- Trijp, R. van. (2008). Argumentstruktur in der Fluid Construction Grammar. In A. Stefanowitsch & K. Fischer (Hrsg.), *Konstruktionsgrammatik II* (Bd. 47, S. 223-246). Tübingen: Stauffenburg.
- Tschichold, C. (2000). *Multi-word units in natural language processing* (Bd. Bd. 6). Hildesheim: Olms.

- Tschichold, C. (2008). A computational lexicography approach to phraseologisms. In S. Granger & F. Meunier (Hrsg.), *Phraseology* (S. 361–376). Amsterdam: John Benjamins.
- Vigliocco, G., Meteyard, L., Andrews, M. & Kousta, S. (2009). Toward a theory of semantic representation. *Language and Cognition*, 1 (2), 219–247.
- Villavicencio, A., Bond, F., Korhonen, A. & McCarthy, D. (2005). Introduction to the special issue on multiword expressions: Having a crack at a hard nut: Special issue on multiword expression. *Computer Speech & Language*, 19 (4), 365–377.
- Wahrig, G. (2007). *Der kleine Wahrig: Wörterbuch der deutschen Sprache: Wörterbuch der deutschen Sprache* (4. Aufl.). München: Bertelsmann.
- Weber, T. (2010). *Lexikon und Grammatik in Interaktion: Lexikalische Kategorisierungsprozesse im Deutschen*. de Gruyter.
- Wipf, S. (2010). *Emotionen im Lebensrückblick alter Menschen: Eine textanalytische Untersuchungen von Verbatimprotokollen auf der Basis des affektiven Diktions der Universität Ulm (ADU)*. Unveröffentlichte Dissertation, Universität Zürich, Psychologisches Institut, Klinische Psychologie, Psychotherapie und Psychoanalyse.
- Wirrer, J. (Hrsg.). (1998). *Phraseologismen in text und kontext*. Bielefeld: Aisthesis Verlag.
- Wolf, M., Horn, A. B., Mehl, M. R., Haug, S., Pennebaker, J. W. & Kordy, H. (2008). Computergestützte quantitative Textanalyse. *Diagnostica*, 54 (2), 85–98.
- Wray, A. (2000). The functions of formulaic language: an integrated model. *Language & Communication*, 20 (1), 1–28.
- Wray, A. (2005). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Wray, A. (2008). *Formulaic language: Pushing the boundaries*. Oxford: Oxford University Press.
- Zeschel, A. (2008). Funktionsverbgefüge als Idiomverbände. In A. Stefanowitsch & K. Fischer (Hrsg.), *Konstruktionsgrammatik II* (Bd. 47, S. 263–278). Tübingen: Stauffenburg.
- Züll, C. & Landmann, J. (2002). *Computerunterstützte Inhaltsanalyse: Literaturbericht zu neueren Anwendungen*.
- Züll, C. & Mohler, P. P. (Hrsg.). (1992). *Textanalyse: Anwendungen der computerunterstützten Inhaltsanalyse*. Opladen: Westdeutscher Verlag.

A Anhang

A.1 Abkürzungsverzeichnis

Tabelle A.1: Abkürzungsverzeichnis

Abkürzung	Beschreibung
AGD	Archiv für Gesprochenes Deutsch
BNC	British National Corpus
CA	conversation analysis / Konversationsanalyse
CHTK	Schweizer-Textkorpus
CPA	Corpus Pattern Analysis
CxG	construction grammar / Konstruktionsgrammatik
DA	Diskursanalyse
DWDS	Digitales Wörterbuch der Deutschen Sprache
FVG	Funktionsverbgefüge ... (support verb constructions,)
GA	Gesprächsanalyse (Gesprächsforschung)
IAL	Interaktionale Linguistik
ICM	Idealized Cognitive Model
IVP	Idiomatisierte verbale Phraseologismen
LSI	Latent Semantic Indexing, Latent Semantic Analysis
MCA	Membership Categorization Analysis
MWE	multi word expression (vgl. auch MWU: multi word unit) = Mehrwortausdruck
MU / MI	Mutual Information: Wahrscheinlichkeitsmass für die Kookurrenz von 2 Lexemen
NLP	Natural Language Processing (z.B.: «NLP- Applikationen»)
NVG	Nominalisierungsverbgefüge
PDEV	Pattern Dictionary of English Verbs
PLI	Phrasal Lexical Items
SKE	sketch engine, ein Werkzeug für Korpusabfragen
VALBU	Valenzwörterbuch deutscher Verben
WSD	Word Sense Disambiguation

A.2 Transkriptionsregeln

Tabelle A.2: Transkriptionsregeln (Mergenthaler, 1992)

Abkürzung	Beschreibung
P:	Patient
T:	Therapeut
	Pseudonym, verschlüsselte Bezeichnung
+	gleichzeitiges Sprechen
///(? :möchte)	unverständlich, Anzahl Schrägstriche = Anzahl Wörter, in Klammern der vermutete Wortlaut
!	Betonung, Hervorhebung
:	Dehnung
?	Frage, steigend oder hoch endende Stimmführung
.	Abgeschlossener Gedanke, auf dem Grundton endende Stimme
,	kurzes Zögern, Gedanke wird jedoch fortgesetzt
;	abgebrochener Gedanke, gefolgt von einem anderen Gedanken
merkwü-	Wortabbruch
Interjektionen:	
hm oder hmhm	Bestätigung
hm?	Frage
hm, oder hm-	Verwunderung
hm,	
hm. oder hm-	Ratlosigkeit
hm.	
hmhm-	Verneinung
Pausen:	
-	2 Sekunden
--	5 Sekunden
---	10 Sekunden
----	15 Sekunden
-----	30 Sekunden
-----	1 Minute
-----	2 Minuten

A.3 Notationsregeln im JAKOB-Lexikon

Im Folgenden werden die formalen Eigenschaften der neuen Lexikonmerkmale beschrieben. Von den originalen OLIF-Kategorien werden nur diejenigen aufgeführt und beschrieben, deren Eigenschaften gegenüber dem OLIF-Konzept abgeändert wurden (McCormick, 2005, siehe auch <http://www.olif.net>). Die Reihenfolge der Merkmale entspricht der Konstruktionsstruktur in der Tabelle (9.7).

canForm

- *Optionale Komponenten* in der `<canForm>` werden nicht erwähnt! D.h. optionale Komponenten sind im Prinzip immer möglich.
- *Variable Komponenten* werden in eckigen Klammern dargestellt.
- Eckige Klammern in der *canForm* werden bei der Suche im Lexikon (man. Suche und Kodierung) weggelassen!

phraseType

Die Werte OLIF-Werte für `<phraseType>` werden ergänzt durch die folgenden Werte: *set phrase, idiom, metaphor, metonymy, somatism*.

synFrame

Die originalen OLIF-Werte werden ersetzt durch die Satzmuster gemäss dem Wörterbuch *Der Kleine Wahrig* (Wahrig, 2007, S. 36f).

natclause

Manuelle Eingabe für das Feld `<natclause>` ist ein einfacher Satz mit den erforderlichen Valenzstellen. Es gibt zwei Kategorien als *Platzhalter* für die variablen Valenzstellen (für die Kasus Nominativ, Akkusativ, Dativ):

- für Personen *jmd, jmdn, jmdm*
- für Sachen/Gegenstände: *etwas, etwasn, etwasm*

Für die Eingabe von Verben mit abtrennebarem Präfix wird ein *weil*-Satz formuliert, damit der Parser das Verb korrekt erkennt (Bsp. «weil jmd abfährt»). Die Konjunktion *weil* wird nach dem Parsingprozess wieder entfernt.

bauplan

Das Merkmal `<bauplan>` wird von der Analyseapplikation auf Grund der Texteingabe im Merkmal `<natclause>` und aus dem Merkmal `<synFrame>` automatisch generiert und ist folgendermassen aufgebaut:

1. *Laufnummer* (Wortnummer), Position dieses Wortes im String. Alle folgenden Positionen sind im String durch ein Leerzeichen abgetrennt.
2. *Grundform des Wortes*. Im Falle der Platzhalter «jmd» und «etwas» wird das Wort unterdrückt (weil variabel) und durch Bindestrich ersetzt; diese Position wird stattdessen durch den semantischen Typ (Pos 6) definiert.
3. *Wortart* (gemäss STTS-Tagset)¹. Auch die Wortart wird bei den Platzhaltern unterdrückt und durch Bindestrich ersetzt.
4. *Link* zur Laufnummer des Root-Elementes. Im Sinne der Dependenzgrammatik ist das Verb die Wurzel der Satzstruktur; Subjekt und Akkusativobjekt besitzen deshalb einen Link zur Position 2 (root).
5. *Syntaktische Informationen* (Konstituente/Satzglied, Kasus, Valenzstelle), z.B.
subj=Subjekt, obja=Akkusativobjekt. Die Satzglieder in der Satzstruktur müssen mit dem vorangestellten Satzmuster übereinstimmen.
6. *Semantischer Typ* gemäss CPA-Ontologie mit aktuell zwei Ausprägungen: Menschliche Akteure (‹Human›) gegenüber Dingen und Sachverhalten (‹Anything›).

Das Merkmal *bauplan* kann anschliessend weiter differenziert werden durch Ersetzen von *Human* und *Anything* durch geeignete Kategorien aus der CPA-Ontologie und durch syntaktische Präzisierungen und Restriktionen, die direkt (ohne Leerschlag) in eckigen Klammern an die entsprechende Position angehängt werden (Bsp. *sg* als Restriktion an der Position 2 *Angst[sg] NN L3 obja n/a* im Lexikoneintrag «Angst haben»).

pattern

Die Notation der Verb-Patterns folgt weitgehend der von Hanks und Pustejovsky vorgeschlagenen (Hanks, 2008a; Pustejovsky et al., 2004). Weil für das JAKOB-Lexikonprojekt und für das Deutsche gewisse Anpassungen nötig sind, ist dieser Standard im folgenden für das JAKOB-Lexikon neu ausformuliert.

- *Verbpatterns* bestehen aus dem Verb im Zentrum und sämtlichen Argumenten (dependenzgrammatisch: Ergänzungen und Angaben, inklusive Subjekt) auf der semantischen, syntaktischen oder lexikalischen Ebene.
- *Doppelte eckige Klammern* umschliessen *semantische Typen*; die Kategorien entstammen der CPA-Ontologie.
- *Optionale semantische Rollen* sind in die semantischen Typen integriert (Bsp. *[[Human=Doctor]]*)

¹Das Stuttgart-Tübingen Tagset (STTS) umfasst 54 Wortartenbezeichner (tags) für deutsche Textkorpora mit Wortartenannotation.

URL: <http://www.sfb441.uni-tuebingen.de/a5/codii/info-stts-de.xhtml> (15.11.2010)

- *Geschweifte Klammern* umschliessen syntaktische Konstituenten, die einen semantischen Typ enthalten können oder ein konkretes Lexem aus dem zugehörigen lexikalischen Set (Pustejovsky et al., 2004).
- *Einfache eckige Klammern* bezeichnen syntaktische Kategorien, nicht konkrete Instanzen. Z.B. bezeichnet [MOD] eine unspezifische Modifikation.
- Runde Klammern bezeichnen optionale Elemente.

construction

Das Merkmal <construction> befindet sich im Versuchsstadium und ist nicht einheitlich festgelegt. Grundsätzlich besteht eine Konstruktion im Lexikon aus allen Merkmalen des Lexikoneintrags. Mit dem Merkmal <construction> könnte die Konstruktion z.B. benannt werden (*Transitivkonstruktion* für eine schematische Konstruktion).

subjField

Im JAKOB-Lexikon zur Verfügung stehende Kategorien sind: *Therapiegespräch, formelles Gespräch, informelles Gespräch, Arztgespräch, Diskussion, Interview, Tagebuch* und *general* als unspezifizierte Gesprächsvariante.

textType

Im JAKOB-Lexikon stehen die folgenden Werte zur Auswahl: *narrative, description, exposition, argumentation, instruction*.

register

Im JAKOB-Lexikon stehen die folgende Kategorien zur Verfügung: *formell, informell, umg (umgangssprachlich), salopp, derb, humoristisch, ironisch, gehoben, baby-talk*.

topic

Dieses Merkmal befindet sich im Versuchsstadium, die Werte können ergänzt werden: *arbeit, befinden, beruf, bewegung, bewusstsein, beziehung, einkaufen, essen, familie, ferien, freizeit, fremd, gefühle, geschäft, gesetz, gesundheit, kleidung, kochen, kunst, körper, liebe, macht, medizin, musik, reisen, religion, restaurant, schule, sex, sozial, spielzeug, sport, streit, tageszeit, technik, theater, therapie, tod, traum, verkehr, wetter*.

Index

- Affektives Diktionär Ulm, 38
agency, 53
Angstwörterbuch, 40
Assoziationsmasse, 115
- Bedeutungskonstitution, 59
Bühnenmetapher, 28
- coercion, 148
Corpus Pattern Analysis, 145
cranberry words, 84
- Dependenz, 121
Dependenztheorie, 120
Diktionär, 38
Diskursmarker, 140
Diskurspattern, 177
Domain, 175
Doppelhelix-Theorie der Sprache, 146
DuELME, 169
Dynamische Bedeutungskonstitution, 281
- emergent grammar, 64
Emergenz, 4, 46, 64, 135
Emotionen, 39
Erzählanalyse JAKOB, 7, 25
 Dimensionen, 32
 Modellierungsfunktionen, 25
- formulaic language, 57, 283
formulation, 24, 136
FrameNet, 163, 197
Funktionsverbgefüge, 105
- Genre, 175
- GermaNet, 162
Gesprochene Sprache, 44, 55, 56, 133
Gesprächsanalyse, 47, 200
Gesprächslinguistik, 47
Gesprächspraktiken, 48
Gesprächstypologie, 174
Grammatikalisierung, 140
- Idiom, 84
idiom breaking, 84
idiom principle, 76
Idiomatisierte verbale
 Phraseologismen, 168
implicitness, 61
Infinitlektionen, 57
Interaktionale Linguistik, 45, 131
Interaktionale Semantik, 138
- JAKOB-Lexikon, 158
- Kernbedeutung, 49, 85, 138
Kognitive Grammatik, cognitive
 grammar, 128
Kollokation, 81, 82, 114
Kommunikative Gattung, 179
Kompositionalitätsprinzip, 59
Konnotation, 169, 185
Konstruktionsgrammatik, 49, 123, 126
- Kontext, 66, 114
Kontextmarker, 167
Kontextualismus, 17
Konversationsanalyse, 50, 52
Kookkurrenz, 115
Korpuslinguistik, 135
Kotext, 114

- lexical choice, 31, 33
- lexical priming, 50, 128
- lexical sets, 147
- Linguistic Inquiry and Word Count, 40
- literal meaning, 35
- meaning construction, 59, 61, 282
- Metonymie, 88
- Mutual Information, 116
- Narrativ, 15–17, 22
- Narrative Identität, 22, 23
- Narrative Psychologie, 17, 23
- Narratives Interview, 18
- Nominalisierungsverbgefüge, 105
- Nominalität, 185
- norms and exploitations, 108, 146, 149, 154, 265, 283
- OLIF, 172
- Ontologie, 147, 181
- open-choice principle, 76
- Oxymoron, 63
- Phraseo-Lex, 169
- Phraseologie, 73
- Phraseologismus, 74, 79
- Positionierung, 22
- preferred argument structure, 184
- Produktivität, 78
- Projektion, 49, 51
- radical construction grammar, 127
- Register, 177
- Situation, 48, 66
- sketch engine, 202
- sketch grammar, 204
- small story, 19
- SPEAKING-Modell, 180
- SPRAC-Modell, 136, 137
- Sprachwandel, 140
- story grammar, 20
- sublanguage, 178
- Synonym, 85
- talk-in-interaction, 47
- text type, 176
- Textanalyse, 37, 38
- Theatermetapher, 28
- therapeutic cycles model, 39
- topic, 179
- Topik, 179
- Transitivkonstruktion, 232
- Unikale Elemente, 84
- Unterspezifizierung, 61, 282
- usage-based model, 124
- UWV-Analysemodell, 165
- Vagheit, 62, 282
- VALBU, 173
- Valenz, 121, 173
- Valenztheorie, 121
- Verbale Phraseologismen, 169
- Verbalität, 185
- Vorgeformtes, 57
- Word Sense Disambiguation, 94
- WordNet, 161, 181
- Wortverbindungen, 166
 - feste, 72
 - freie, 72
- Wortverbindungsfelder, 166
- Wortverbindungsmuster, 166, 167

Konstruktionsindex

- beruhigen, 148
- bringen
 - auf die Palme bringen, 77
 - den Stein ins Rollen bringen, 78, 142, 149
- commercial event, 62
- Dampf
 - Dampf ablassen, 168, 266
 - Dampf aufladen, 266
 - Dampf machen, 266
 - seinen Dampf loswerden, 149, 266
- Dampf ablassen, 86
- Dampf aufbauen, 86, 283
- Dank sagen, 79
- das fünfte Rad am Wagen sein, 79
- das Tischtuch vom Tisch niesen (Fred sneezed the napkin off the table), 63
- dastehen wie der Esel am Berg, 241
- den Vogel abschiessen, 169
- die Achseln zucken, 79
- eine Schwalbe macht noch keinen Sommer, 129
- einen Bären aufbinden, 78, 92
- einen Streit vom Zaun brechen, 79, 169
- fressen
 - einen Narren an jmdm fressen, 77
 - Kreide fressen, 77
- geben
 - den Ausschlag geben, 84, 226
 - die Hand geben, 131
 - jmdm einen Korb geben, 79
 - jmdm Saures geben, 143
- im Regen stehen, 97
- ins Gehege zu geraten, 236
- ins Gras beißen, 84, 169
- ins Gärtchen trampen, 237, 243
- jmdn Lügen strafen, 84
- Karren
 - an den Karren fahren, 234
 - den Karren ziehen, 240
- können nicht mehr, 28
- leisten, 104
 - einen Beitrag leisten, 210
 - Hilfe leisten, 104
 - leisten als Funktionsverb, 209
- Maulaffen feilhalten, 84
- seine Flossen hinstrecken, 28
- sich die Zähne putzen, 72
- sich zu Tode langweilen, 34
- sicher sein, 62
- Sorge tragen, 244
- to cry for (the) moon, 149
- Transitivkonstruktion, 127
- verjagen, 258, 263
- verstehen, 133
- Zeug, 249
 - ein Zeug machen, 251
 - im Zeug sein, 251

Lebenslauf

Name:	Marc Luder
Adresse:	In Grosswiesen 16, 8044 Gockhausen
Tel.Nr:	043 355 03 60
Email:	m.luder@psychologie.uzh.ch
Geb.datum:	13. März 1948
Zivilstand:	verheiratet 3 Kinder (Jg. 1973, 1975, 1999)
Schulbildung	1955–1965 Primar- und Sekundarschulen in Bern und Thun 1965–1968 Gymnasium in Thun, Matura 1969 Rekrutenschule
Berufliche Tätigkeiten:	1970–1972 Schauspielschule in Zürich (Schauspielakademie) 1972–1974 Bewirtschaftung eines Bauernhofes in Steg (Tösstal) 1974–1984 Pacht eines Landwirtschaftsbetriebes in Weggis (LU), selbstständiger Landwirt 1979–1983 Jazzschule Luzern (Arrangement, Piano), Aufbau einer Big Band im Raume Luzern 1980–1984 Aufbau eines Kurszentrums auf dem Bauernhof in Weggis. Angebot von Tagungsräumen und Leitung von Musikimprovisationskursen (Jazz) 1985–1996 Pacht des Hotels «Rütliblick» in Morschach (Zentral-schweiz) und Betrieb als Kulturzentrum und Hotel. Geschäftsführer der Genossenschaft bis Ende 1995. Arbeit als Kursleiter und Administrator. 1993–1999 Psychologiestudium an der Universität Zürich (Klinische Psychologie), Nebenfächer Psychopathologie und Computerlinguistik, Lizentiat. 1996–2000 Softwareentwicklung für eine psychologische Praxis. Softwareentwickler bei der Firma Greenshare AG in Wetzikon. ab Mai 2000 Wissenschaftlicher Mitarbeiter am Psychologischen Institut der Universität Zürich, Klinische Psychologie. ab 2001 Betreuung eines ICT-Projektes (Informations- und Kommunikationstechnologie) am Psychologischen Institut der Universität Zürich: Entwicklung eines internetbasierten Textanalyseystems (Erzählanalyse JAKOB). ab 2002 Wissenschaftlicher Mitarbeiter Informatikdienste der Universität Zürich: medizinische E-Learning Plattform (VAM) und OLAT, 50% Pensum. ab 2008 Wissenschaftlicher Mitarbeiter am Psychologischen Institut der Universität Zürich, Klinische Psychologie, Psychotherapie und Psychoanalyse (Lehrstuhl B. Boothe). Dis-sertationsprojekt.

